A finite information KAM theorem

Piotr Kamieński

Contents

1 Introduction

The goal of this paper is to present a methodology for proving KAM-type theorems without the assumption of the classical diophantine condition. Instead of it we base our reasoning only on finite approximation of the associated frequency ω and an arithmetic *Khintchine - Lévy condition*, which we prove to be generic in the sense of measure theory. In other words we give (a lower bound for) the admissible perturbation threshold which depends only on the approximation of the frequency ω appearing in the problem provided that ω does not belong to a set of small measure.

The proofs are constructive in the sense that for concrete systems one can retrieve numerical values of the quantities involved. We emphasize that all the constants appearing in all our theorems and lemmas (whether they concern diophantine approximation, probability or dynamics) are explicitly computed and of numerically reasonable size. This way our techniques can be applied in computer assisted proofs.

We demonstrate our approach using one of the least complicated small divisors problems - the problem of analytic rectification of a small perturbation of a constant vector field on the two-dimensional torus. In our method we follow the Kolmogorov-Newton iterative scheme. Specifically we ask under what conditions we can transform the equation

$$\begin{cases} \dot{x} = 1 + a^*(x,y) \\ \dot{y} = \omega + a^{**}(x,y) \end{cases} \tag{1}$$

(here $a = (a^*, a^{**})$ is a small function and $x, y \in \mathbb{S}^1$) to the equation

$$\begin{cases} \dot{X} = 1 \\ \dot{Y} = \omega \end{cases} \tag{2}$$

by means of an analytic, close to identity change of variables $(x,y) \mapsto (X,Y)$. In our method we follow the Kolmogorov-Newton iterative scheme. The key ingredient is an estimate on the size of the solution of the homological equation, which we obtain by a careful analysis of the small divisors.

Assuming an arithmetic condition on ω is, of course, inevitable as the denominators appearing in the problem (in our case $|q\omega - p|$ for integers p and q) can be arbitrarily close to zero. However, for a majority of indices (p,q) the size of the denominators can be controlled. The Khintchine - Lévy arithmetic condition, which we introduce deals with the remaining ones, which are related to p_n/q_n - the continued fraction convergents to ω - by assuming a certain growth rate of the sequence q_n.

Historically KAM theory originates from papers of Kolmogorov, Arnol'd and Moser ([17, 1, 2, 22]), who showed that invariant tori which appear naturally in integrable Hamiltonian systems persist sufficiently small perturbations of these systems under the assumption of two conditions: the geometric twist condition and a number-theoretic diophantine condition on the frequencies of the investigated torus. Fundamental questions which arise in the context of these theorems involve the maximal size of perturbation admissible by their assumptions, its dependence on the frequencies of the torus and the structure (and in particular measure) of the set of these frequencies, for which the results can be applied.

The aforementioned seminal papers did not, however, answer these questions of quantitative nature, instead their authors focused mainly on providing conditions, which ensure the validity of the theorems for undetermined, sufficiently small perturbation sizes. Since then only a fraction of papers on KAM theory concentrated on its quantitative aspect and only partial results in this area are available. First progress has been done by Rüssmann ([25, 26]), who improved the estimates on the solution of the homological equation (in the case of diophantine frequency vectors), which resulted in a better lower bound of the admissible perturbation size. Later Herman in [14] gave a lower estimate in the case of a very special rotation number and a very special transformation: the golden mean circle in the Chirikov standard map. This has been improved - using computer assisted methods - by de la Llave and Rana and Celletti and Chierchia ([13, 12, 6, 7]). We should also mention the paper [9] in which from this perspective the Siegel center problem has been analyzed as well as [11] in which the authors develop methods applicable to any Hamiltonian system, not necessarily a close to integrable one.

The aforementioned Siegel problem ([29]) has also been investigated by Yoccoz, Marmi, Buff and Chéritat. Yoccoz proved that in the quadratic polynomial case the necessary and sufficient condition for linearizing a perturbation of a rotation in the complex plane is the convergence of the Brjuno function of the associated frequency. Buff and Cheritat went further proving there is a close connection between the Brjuno function and the radius of convergence of the change of variables in the Siegel problem. In the context of Hamiltonian systems the Brjuno function has been studied by Rüssmann ([27]) and measure-theoretic aspects of KAM theory have been investigated by Pöschel ([24]) and since his results no new methods concerning this problem have been invented.

A survey of history and results of the theory of small divisors may be found in [10] and [5] in the context of Hamiltonian systems and in [30, 8, 31, 19] and [20] in the context of the Siegel problem and Brjuno function.

In our paper we do not deal with the classical diophantine condition, instead we concentrate on the continued fraction expansion of the frequency involved. We do not, however, look at ω pointwise, but globally - we use the statistical properties of the continued fraction transformation. The Khintchine-Lévy condition is motivated by the theorem on the Khintchine-Lévy constant (theorem 2.1, [18]). This result allows us to expect that even though the growth rate of the sequence (q_n) and other similar sequences can be arbitrarily fast, the typical behavior - in the sense of measure theory - is exponential growth. The original paper of Lévy, however, concerns only pointwise a.e. convergence to a constant of the sequence $\sqrt[n]{q_n}$ and it lacks quantitative nature. This has been improved in i.e. [15, 23] and [21] to a central limit theorem, law of the iterated logarithm and a Berry-Esséen type bound (a large deviations result). In our paper we use large deviations theorems specifically tailored for our needs to obtain estimates of the measure of the set of numbers verifying the Khintchine-Lévy condition.

The paper is organized as follows. In section 2 we state our results after introducing all the necessary tools. In section 3 we prove a lemma on diophantine approximation, which allows us to deal with the "medium-size" denominators. Using these results in section 4 we analyze our main object of investigation - the homological equation. Finally section 5 contains estimates of the measure of the Khintchine-Lévy set.

2 Tools, result and consequences

We will be working with irrational numbers, but since we would like them to constitute a probability space we will restrict our attention to the set $X := (0, 1) \setminus \mathbb{Q}$. For a set $A \subset X$ we denote its complement $X \setminus A$ as A^c. The σ-algebra of sets in use will be the family of Lebesgue measurable sets and by λ we denote the Lebesgue measure on X. We will also need the Gauss measure γ given by

$$\gamma(B) := \frac{1}{\log 2} \int_B \frac{d\lambda(x)}{1+x}. \tag{3}$$

These two measures are absolutely continuous with respect to one another. The term "almost all" will mean "λ-almost all" when used with respect to elements of X and "all except for a finite number" when refered to sets of indices, in particular the integers or natural numbers. The symbol $\mathbb{E}_\mu$ will denote the expected value of a random variable with respect to a measure μ.

Let Y be a random variable defined on a probability space $(\mathcal{Y}, \Sigma, \mu)$. Denote by F_Y its distribution function, i.e.

$$F_Y(x) = \mu(Y < x) = \mu(Y^{-1}((-\infty, x))) \tag{4}$$

and by f_Y the characteristic function of Y:

$$f_Y(t) = \mathbb{E}_\mu \exp(itY). \tag{5}$$

It is a well known fact that the moments of Y and derivatives of f_Y at zero are related by

$$\mathbb{E}_\mu Y^\nu = \frac{1}{i^\nu}\frac{d^\nu}{dt^\nu} f_Y(t)\Big|_{t=0} \tag{6}$$

We define the *cumulants* of Y (with respect to μ) as

$$\Gamma_k(Y) = \frac{1}{i^k}\frac{d^k}{dt^k}\left(\log f_Y(t)\right)\Big|_{t=0}. \tag{7}$$

For a fixed real number ω we denote by $[a_0; a_1, a_2, \ldots]$ its continued fraction expansion:

$$\omega = [a_0; a_1, a_2, \ldots] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots}}}}, \tag{8}$$

Here $a_j = a_j(\omega)$ are positive integers, we will refer to them as *partial quotients of* ω.

On $[0, 1)$, where $a_0 = 0$ we will often write $[a_1, a_2, \ldots]$ instead of $[0; a_1, a_2, \ldots]$. There the shift on the continued fraction expansion is given by the *Gauss map* $G : X \mapsto X$:

$$G(\omega) = G([a_1, a_2, \ldots]) = [a_2, a_3, \ldots] = \left\{\frac{1}{\omega}\right\}. \tag{9}$$

Here $\{\cdot\}$ denotes the fractional part (we will also use $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$ for floor and ceiling functions, respectively). The Gauss map has two properties, which will be essential for us: it leaves the Gauss measure γ invariant and it is ergodic with respect to that measure.

By $\frac{p_n}{q_n}$, where p_n and q_n are coprime integers, we denote the number

$$\frac{p_n}{q_n} = [a_0; a_1, \ldots, a_n] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}, \tag{10}$$

which we call the *n-th convergent to* ω. We will also need the *complete quotiens* x_n defined by:

$$x_n(\omega) = [a_n; a_{n+1}, a_{n+2}, \ldots] = \frac{1}{G^{n-1}(\omega - a_0)} \tag{11}$$

as well as the *approximation errors* η_n:

$$\eta_n = |q_n\omega - p_n|. \tag{12}$$

The following lemma collects some textbook results on continued fractions. We will often use them silently furtheron.

Lemma 2.1. *For any irrational* ω *the following relations are satisfied:*

$$q_n = a_n q_{n-1} + q_{n-2}, \qquad n \geqslant 0; \qquad q_{-2} = 1, q_{-1} = 0, \tag{13}$$

$$p_n = a_n p_{n-1} + p_{n-2}, \qquad n \geqslant 0; \qquad p_{-2} = 0, p_{-1} = 1, \tag{14}$$

$$\eta_n = \eta_{n-2} - a_n\eta_{n-1}, \quad n \geqslant 0; \quad \eta_{-2} = \omega, \eta_{-1} = 1, \tag{15}$$

$$\eta_n = (-1)^n(q_n\omega - p_n), \quad n \geqslant -2; \tag{16}$$

$$x_n = \frac{\eta_{n-2}}{\eta_{n-1}}, \quad n \geqslant 0; \tag{17}$$

$$x_1x_2\dots x_n = \eta_{n-1}^{-1}, \quad n \geqslant 1; \tag{18}$$

$$\frac{1}{q_{n+1}+q_n} < |q_n\omega - p_n| < \frac{1}{q_{n+1}}, \quad n \geqslant 0. \tag{19}$$

The convergents p_n/q_n will be important for us, since they are the best rational approximations to ω and they give rise to the smallest of the small divisors.

Lemma 2.2. *If $\frac{p_n}{q_n}$ is a convergent to an irrational number ω the following inequality holds for all p and all $q < q_n$:*

$$|q\omega - p| > |q_n\omega - p_n|. \tag{20}$$

The numerators and denominators of the convergents, however, are not the only indices, which are critical from the point of view of KAM theory. Among them are also the ones introduced in the following definition, which will be justified in section 3.

Definition 2.1. *Let ω be an irrational number. We define the sets of*

- initial denominators of ω
$$ID = \{1, \dots, a_1\}, \tag{21}$$
- true denominators of ω
$$TD = \{q_n, \ \textit{where } n \geqslant 2\}, \tag{22}$$
- semidenominators of ω
$$SD = \{aq_{n-1} + q_{n-2}, \ \textit{where } a \in \{1, \dots, a_n - 1\} \textit{ and } n \geqslant 2\}, \tag{23}$$
- cosemidenominators of ω
$$CSD = \{aq_{n-1}, \ \textit{where } a \in \{2, \dots, a_n\} \textit{ and } n \geqslant 2\}. \tag{24}$$

The sets ID, TD, SD and CSD are pairwise disjoint.

By ℓ we understand the logarithm of the Khintchine-Lévy constant (or the Khintchine-Lévy constant itself, depending on the convention), that is $\ell := \frac{\pi^2}{12\log 2}$. Its importance is demonstrated in the following

Theorem 2.1 (Khintchine-Lévy theorem, [18])**.** *The sequence $\frac{1}{n}\log q_n(\omega)$ tends to ℓ for almost every $\omega \in X$.*

In view of the above result it is reasonable to assume that on a big set of numbers ω the growth rate of $q_n(\omega)$ could be bounded from above by a sequence only slightly faster than $e^{\ell n}$ and similarly from below by a slightly slower sequence. This justifies the following definition and will be made precise in section 5.

Definition 2.2. *We say that an irrational number ω satisfies the* upper Khintchine-Lévy condition *with constants $T > 0$ and $N \in \mathbb{N}$ (or simply is (T, N)-upper-Khintchine-Lévy) if the following inequality holds for all $n \geqslant N$:*

$$q_n(\omega) \leqslant e^{(\ell+T)n}. \tag{25}$$

We denote the set of (T,N)-upper-Khintchine-Lévy numbers by $KL^+(T,N)$. Analogously we define the lower Khintchine-Lévy *set by*

$$KL^-(T,N) = \left\{\omega \in X : \frac{1}{2}e^{(\ell - T)n} \leqslant q_n(\omega) \text{ for all } n \geqslant N\right\}. \tag{26}$$

By $KL(T_-,T_+,N)$ we will understand the intersection of $KL^+(T_+,N)$ and $KL^-(T_-,N)$ and we will write $KL(T,N)$ for $KL(T,T,N)$.

The role of the $1/2$ in the definition of $KL^-(T,N)$ is purely technical and will become clear in section 5. Also, for a given natural number n, we denote by $KL_n^+(T)$ the set $\{\omega \in X : q_n(\omega) \leqslant e^{(\ell+T)n}\}$ and similarly for KL^- (with the $1/2$ factor) and KL. This way

$$KL^\circ(T,N) = \bigcap_{n=N}^{\infty} KL_n^\circ(T) \tag{27}$$

with $\circ \in \{+,-,\ \}$.

Remark 2.3. *In the case when $T_- > \ell - \log\frac{1+\sqrt{5}}{2} \approx 0.705$ the set $KL^-(T_-,N)^c$ is empty for all N, as the continued fraction denominators are bounded from below by the Fibonacci sequence for all numbers ω.*

The main object of study of the remaining part of this paper will be the following linear PDE, called the *homological equation*:

$$(\partial_x + \omega\partial_y)h(x,y) = a(x,y), \tag{HOM}$$

where h is the unknown, a is given and both of them are real-valued maps with domain $\mathbb{R}^2$, which are 2π-periodic in both x and y directions. This periodicity allows us (after imposing some mild regularity conditions) to write a and h as Fourier series:

$$a(x,y) = \sum_{l\in\mathbb{Z}^2} a_l e^{i(l_1x+l_2y)}, \tag{28}$$

$$h(x,y) = \sum_{l\in\mathbb{Z}^2} h_l e^{i(l_1x+l_2y)}, \tag{29}$$

on which we impose a reality condition $\bar{\cdot}_l = \cdot_{-l}$. Using expansions (28) and (29) it is not hard to see that h_l's need to satisfy

$$h_l = \frac{a_l}{i(l_1+\omega l_2)}. \tag{30}$$

First we have to make sure we are not dividing by zero, which is the case only if $l_1 = l_2 = 0$ when ω is irrational. To achieve this we must only work with a's for which a_0, their mean value, is equal to zero. However, this restriction leaves us with the possibility to choose the value of h_0, for simplicity we will always choose $h_0 = 0$.

To be more specific, we will be working in an analytic setting. First let us define the (scale of) space(s) which we will make use of. Let ρ be a positive real number. We set

$$\Pi(\rho) := \{z \in \mathbb{C} : |\mathrm{Im}\, z| \leqslant \rho\}, \tag{31}$$

and

$$\mathcal{P}_\rho := \{f : \Pi(\rho)^2 \mapsto \mathbb{C} : f \text{ is analytic in } \mathrm{int}\,\Pi(\rho)^2\text{, continuous, } 2\pi\text{-periodic in both directions and } f(\mathbb{R}^2) \subset \mathbb{R}\}. \tag{32}$$

On $\mathcal{P}_\rho$ we use the standard sup-norm:

$$||f||_\rho := \sup_{x,y\in\Pi(\rho)} |f(x,y)|. \tag{33}$$

For a multiindex $(l_1,l_2) = l \in \mathbb{Z}^2$ we will denote by $|l|$ its ℓ^1 norm, namely $|l| = |l_1| + |l_2|$. The symbol S_*, where S is a set of indices over which some sum is taken, means simply $S \setminus \{0\}$.

The main result of this paper is the following

Theorem 2.2. *Suppose $\omega \in KL(T_-, T_+, N)$ for some $T_-, T_+ > 0$ and $N \in \mathbb{N}_*$ and that a is a zero-average function in $\mathcal{P}_\rho$ for some $\rho > 0$. Then the homological equation* (HOM) *has a unique zero-average solution h, which belongs to $\mathcal{P}_{\rho-\delta}$ for any $0 < \delta < \rho$ and*

$$||h||_{\rho-\delta}||a||_\rho^{-1} \leqslant \frac{9\pi}{2}C\delta^{-1} + 2S + 2S' + e^\delta e^{\ell+T_+}\left[D_1(\delta) + D_2(\delta) + D_3(\delta)\right]\delta^{-L} + \mathsf{brj}_\omega^{(N)} - (S_1 + S_2 + S_3), \tag{34}$$

where

$$C = \max\left\{2, \frac{1}{\lceil\omega\rceil - \omega}\right\}, \tag{35}$$

$$L = \frac{\ell + T_+}{\ell - T_-}, \tag{36}$$

$$S = \sum_{q=1}^{a_1} \frac{e^{-(q+qa_0+1)\delta}}{|q\omega - qa_0 - 1|}, \tag{37}$$

$$S' = \sum_{q=1}^{a_1} \frac{e^{-q(1+a_0)\delta}}{q|\omega - a_0|}, \tag{38}$$

$$D_1(\delta) = 2^L\Gamma(L) + (2L)^L e^{-1/4L}, \tag{39}$$

$$D_2(\delta) = \frac{\log 2 + \ell + T_+}{\log 2}\left[\Gamma(L)\cdot 2^L\left(1 + e^{-(\ell-T_-)}\right)^L + (2L)^L e^{-\left(1+e^{-(\ell-T_-)}\right)/4L}\right] + \tag{40}$$

$$+ \frac{T_+ - T_-}{\log 2}\left[\frac{2^L}{(\ell-T_-)^2\left(1+e^{-(\ell+T_+)}\right)^L}\left(\Gamma'(L) + \Gamma(L)\log\frac{2}{1+e^{-(\ell+T_+)}} + \Gamma(L)\log\frac{1}{(1+\omega)\delta}\right)\right], \tag{41}$$

$$D_3(\delta) = \Gamma(L)\log 2 + L^L e^{-1/2L}\log 2 + \frac{T_+ - T_-}{(\ell-T_-)^2}\Gamma'(L) + \frac{T_+ - T_-}{(\ell-T_-)^2}\Gamma(L)\log\frac{1}{(1+\omega)\delta}, \tag{42}$$

$$S_1 = \sum_{n=1}^{N-1} e^{(\ell+T_+)n}\exp\left(-\frac{1}{2}e^{(\ell-T_-)n}(1+\omega)\delta\right), \tag{43}$$

$$S_2 = \frac{1}{\log 2}\sum_{n=1}^{N-1} e^{(\ell+T_+)n}\exp\left(-\frac{1}{2}(1+\omega\delta\left(1+e^{-(\ell-T_-)}\right)e^{(\ell-T_-)n}\right)\cdot\left[\log 2 + \ell + T_+ + (T_+ - T_-)n\right], \tag{44}$$

$$S_3 = \sum_{n=1}^{N-1} e^{(\ell+T_+)(n+1)}\exp\left(-(1+\omega)\delta e^{(\ell-T_-)n}\right)\left[\log 2 + (T_+ - T_-)n\right], \tag{45}$$

$\mathsf{brj}_\omega^{(N)}$ *is given in definition 4.3 and Γ is the Euler Γ function.*

Observe that the bound in theorem 2.2 is $O(\delta^{-L}\log\delta^{-1})$. This allows us to perform the Kolmogorov-Newton iterative scheme - we give a sketch of the procedure below.

We consider a perturbation of the constant vector field $(1, \omega)$ on the 2-torus:

$$\begin{cases} \dot{x} = 1 + a^*(x,y) \\ \dot{y} = \omega + a^{**}(x,y) \end{cases} , \tag{46}$$

where a^* and a^{**} are 2π-periodic in x and y and they extend to analytic functions on $\Pi(\rho^*) \times \Pi(\rho^{**})$. Additionally assume that their averages are equal to 0 and div $a = 0$ (where $a = (a^*, a^{**})$), which makes our perturbation Hamiltonian. We will be looking for a measure-preserving (symplectic) coordinate change $(x, y) \mapsto (X, Y)$, where

$$\begin{cases} X = x + h^*(x,y) \\ Y = y + h^{**}(x,y) \end{cases} \tag{47}$$

and

$$\begin{cases} X = x + g^*(x,Y) \\ y = Y + g^{**}(x,Y) \end{cases} \tag{48}$$

We will use the second, implicit definition more often by analogy with the generating function method for defining symplectic transformations. Observe also that to require that $(x, y) \mapsto (X, Y)$ is measure preserving is the same as requiring that div $g = 0$ (with $g = (g^*, g^{**})$). Let us now calculate the perturbation in these new coordinates. We want to obtain

$$\begin{cases} \dot{X} = 1 + \hat{a}^*(X,Y) \\ \dot{Y} = \omega + \hat{a}^{**}(X,Y) \end{cases} . \tag{49}$$

with $||\hat{a}|| = O(||a||^2)$. We have

$$\begin{cases} \dot{X} = \dot{x} + \dot{x}\partial_x g^*(x,Y) + \dot{Y}\partial_Y g^*(x,Y) \\ \dot{y} = \dot{Y} + \dot{x}\partial_x g^{**}(x,Y) + \dot{Y}\partial_Y g^{**}(x,Y) \end{cases} \tag{50}$$

$$\begin{cases} 1 + \hat{a}^*(X,Y) = 1 + a^*(x,y) + (1 + a^*(x,y))\partial_x g^*(x,Y) + (\omega + \hat{a}^{**}(X,Y))\partial_Y g^*(x,Y) \\ \omega + a^{**}(x,y) = \omega + \hat{a}^{**}(X,Y) + (1 + a^*(x,y))\partial_x g^{**}(x,Y) + (\omega + \hat{a}^{**}(X,Y))\partial_Y g^{**}(x,Y) \end{cases} \tag{51}$$

$$\begin{cases} \hat{a}^*(X,Y) = a^*(x,y) + (\partial_x + \omega\partial_Y)g^*(x,Y) + a^*(x,y)\partial_x g^*(x,Y) + \hat{a}^{**}(X,Y)\partial_Y g^*(x,Y) \\ a^{**}(x,y) = \hat{a}^{**}(X,Y) + (\partial_x + \omega\partial_Y)g^{**}(x,Y) + a^*(x,y)\partial_x g^{**}(x,Y) + \hat{a}^{**}(X,Y)\partial_Y g^{**}(x,Y) \end{cases} \tag{52}$$

$$\begin{cases} \hat{a}^*(X,Y) = a^*(x,y) - a^*(x,Y) + a^*(x,Y) + (\partial_x + \omega\partial_Y)g^*(x,Y) + a^*(x,y)\partial_x g^*(x,Y) + \hat{a}^{**}(X,Y)\partial_Y g^*(x,Y) \\ -\hat{a}^{**}(X,Y) = \dfrac{-a^{**}(x,y) + a^{**}(x,Y) - a^{**}(x,Y) + (\partial_x + \omega\partial_Y)g^{**}(x,Y) + a^*(x,y)\partial_x g^{**}(x,Y)}{1 + \partial_Y g^*(x,Y)} \end{cases} \tag{53}$$

Now assume that g^* and g^{**} solve the homological equations with initial data $-a^*$ and a^{**}, respectively. Equation (53) becomes

$$\begin{cases} \hat{a}^*(X,Y) = a^*(x,y) - a^*(x,Y) + a^*(x,y)\partial_x g^*(x,Y) + \hat{a}^{**}(X,Y)\partial_Y g^*(x,Y) \\ -\hat{a}^{**}(X,Y) = [-a^{**}(x,y) + a^{**}(x,Y) + a^*(x,y)\partial_x g^{**}(x,Y)] \cdot (1 + \partial_Y g^*(x,Y))^{-1} \end{cases} \tag{54}$$

$$\begin{cases} \hat{a}^*(X,Y) = a^*(x, Y + g^{**}(x,Y)) - a^*(x,Y) + a^*(x, Y + g^{**}(x,Y))\partial_x g^*(x,Y) + \hat{a}^{**}(X,Y)\partial_Y g^*(x,Y) \\ -\hat{a}^{**}(X,Y) = [-a^{**}(x, Y + g^{**}(x,Y)) + a^{**}(x,Y) + a^*(x, Y + g^{**}(x,Y))\partial_x g^{**}(x,Y)] \cdot (1 + \partial_Y g^*(x,Y))^{-1} \end{cases} \tag{55}$$

For a function $b : \Pi(R^*) \times \Pi(R^{**}) \ni (\xi, \eta) \mapsto b(\xi, \eta) \in \mathbb{C}$ define

$$||b||_{R^*,R^{**}} = \sup_{(\xi,\eta)\in\Pi(R^*)\times\Pi(R^{**})} |b(\xi, \eta)|. \tag{56}$$

We have the Cauchy estimates:

$$||\partial_\xi b||_{R^*-\mu,R^{**}} \leqslant \mu^{-1}||b||_{R^*,R^{**}}, \tag{57}$$

$$||\partial_\eta b||_{R^*,R^{**}-\nu} \leqslant \nu^{-1}||b||_{R^*,R^{**}}. \tag{58}$$

Assume additionally that the following estimates hold for all δ and some function Γ:

$$||v||_{R^*-\delta,R^{**}-\delta} \leqslant \Gamma(\delta)||b||_{R^*,R^{**}}, \tag{59}$$

where v is the solution of the homological equation with initial data b.

For a moment fix a $\delta > 0$ and assume $R^* < \rho^* - \delta$ and $R^{**} < \rho^{**} - \delta$. Define

$$\beta^* = \Gamma(\delta)||a^*||_{\rho^*,\rho^{**}} \tag{60}$$

and

$$\beta^{**} = \Gamma(\delta)||a^{**}||_{\rho^*,\rho^{**}}. \tag{61}$$

Observe, that if $(x, Y) \in \Pi(R^*) \times \Pi(R^{**})$ then

$$|X - x| \leqslant ||g^*||_{R^*,R^{**}} \leqslant \beta^* \tag{62}$$

and similarily

$$|y - Y| \leqslant ||g^{**}||_{R^*,R^{**}} \leqslant \beta^{**}. \tag{63}$$

Thus if $X \in \Pi(r - \beta^*)$ then $x \in \Pi(r)$, whenever $r - \beta^* < \rho^* - \delta$. This in turn tells us that the supremum of some quantity depending on X and Y taken over $(X, Y) \in \Pi(R^* - \beta^*) \times \Pi(R^{**})$ is smaller than the supremum of this quantity written in terms of x and Y, which is taken over $(x, Y) \in \Pi(R^*) \times \Pi(R^{**})$ (provided that $R^* - \beta^* < \rho^* - \delta$ and $R^{**} < \rho^{**} - \delta$).

We now make an attempt to estimate the supremum of $-\hat{a}^{**}(X, Y)$ on the strip

$$\Pi(\rho^* - \delta - \beta^*, \rho^{**} - \max\{\delta + \mu, \delta + \nu, \nu + \beta^{**}\}). \tag{64}$$

In order to shorten notation we write

$$\sup_{\substack{\xi:R^*\\ \eta:R^{**}}} \quad \text{for} \quad \sup_{(\xi,\eta)\in\Pi(R^*)\times\Pi(R^{**})}. \tag{65}$$

We have

$$\sup_{\substack{X:\rho^*-\delta-\beta^*\\ Y:\rho^{**}-\max\{\delta+\mu,\delta+\nu,\nu+\beta^{**}\}}} |\hat{a}^{**}(X, Y)| \leqslant (I^{**} + II^{**}) \cdot III^{**}, \tag{66}$$

where

$$I^{**} := \sup_{\substack{x:\rho^*-\delta\\ Y:\rho^{**}-\max\{\delta+\mu,\delta+\nu,\nu+\beta^{**}\}}} |a^{**}(x, Y + g^{**}(x, Y)) - a^{**}(x, Y)|, \tag{67}$$

$$II^{**} := \sup_{\substack{x:\rho^*-\delta\\ Y:\rho^{**}-\max\{\delta+\mu,\delta+\nu,\nu+\beta^{**}\}}} |a^*(x, Y + g^{**}(x, Y)) \cdot \partial_x g^{**}(x, Y)|, \tag{68}$$

$$III^{**} := \sup_{\substack{x:\rho^*-\delta \\ Y:\rho^{**}-\max\{\delta+\mu,\delta+\nu,\nu+\beta^{**}\}}} |1+\partial_Y g^*(x,Y)|^{-1}. \tag{69}$$

We estimate I^{**}, II^{**} and III^{**} separately.

$$I^{**} \leqslant \sup_{\substack{x:\rho^*-\delta \\ Y:\rho^{**}-\max\{\delta+\mu,\delta+\nu,\nu+\beta^{**}\}}} \left(|g^{**}(x,Y)| \cdot \sup_{\eta\in[0,g^{**}(x,Y)]} |\partial_Y a^{**}(x,Y+\eta)| \right) \leqslant \tag{70}$$

$$\leqslant \left(\sup_{\substack{x:\rho^*-\delta \\ Y:\rho^{**}-\max\{\delta+\mu,\delta+\nu,\nu+\beta^{**}\}}} |g^{**}(x,Y)| \right) \cdot \left(\sup_{\substack{x:\rho^*-\delta \\ Y:\rho^{**}-\max\{\delta+\mu-\beta^{**},\delta+\nu-\beta^{**},\nu\}}} |\partial_Y a^{**}(x,Y)| \right) \leqslant \tag{71}$$

$$\leqslant \Gamma(\delta)||a^{**}||_{\rho^*,\rho^{**}} \cdot \nu^{-1}||a^{**}||_{\rho^*,\rho^{**}} = \nu^{-1}\Gamma(\delta)||a^{**}||^2_{\rho^*,\rho^{**}}, \tag{72}$$

$$II^{**} \leqslant \left(\sup_{\substack{x:\rho^*-\delta \\ Y:\rho^{**}-\max\{\delta+\mu-\beta^{**},\delta+\nu-\beta^{**},\nu\}}} |a^*(x,Y)| \right) \cdot \mu^{-1} \left(\sup_{\substack{x:\rho^*-\delta \\ Y:\rho^{**}-\max\{\delta,\delta+\nu-\mu,\nu+\beta^{**}-\mu\}}} |g^{**}(x,Y)| \right) \leqslant \tag{73}$$

$$\leqslant \mu^{-1}\Gamma(\delta)||a^*||_{\rho^*,\rho^{**}}||a^{**}||_{\rho^*,\rho^{**}}, \tag{74}$$

$$III^{**} \leqslant \sup_{\substack{x:\rho^*-\delta \\ Y:\rho^{**}-\max\{\delta+\mu,\delta+\nu,\nu+\beta^{**}\}}} (1-|\partial_Y g^*(x,Y)|)^{-1} \leqslant \left(1-\nu^{-1}\Gamma(\delta)||a^*||_{\rho^*,\rho^{**}}\right)^{-1} \tag{75}$$

Using the above we can now estimate the supremum of $\hat{a}^*(X,Y)$ over the same set:

$$\sup_{\substack{X:\rho^*-\delta-\beta^* \\ Y:\rho^{**}-\max\{\delta+\mu,\delta+\nu,\nu+\beta^{**}\}}} |\hat{a}^*(X,Y)| \leqslant I^* + II^* + III^*, \tag{76}$$

where

$$I^* := \left(\sup_{\substack{X:\rho^*-\delta-\beta^* \\ Y:\rho^{**}-\max\{\delta+\mu,\delta+\nu,\nu+\beta^{**}\}}} |\hat{a}^{**}(X,Y)| \right) \cdot \left(\sup_{\substack{x:\rho^*-\delta \\ Y:\rho^{**}-\max\{\delta+\mu,\delta+\nu,\nu+\beta^{**}\}}} |\partial_Y g^*(x,Y)| \right), \tag{77}$$

$$II^* := \sup_{\substack{x:\rho^*-\delta \\ Y:\rho^{**}-\max\{\delta+\mu,\delta+\nu,\nu+\beta^{**}\}}} |a^*(x,Y+g^{**}(x,Y)) - a^*(x,Y)|, \tag{78}$$

$$III^* := \sup_{\substack{x:\rho^*-\delta \\ Y:\rho^{**}-\max\{\delta+\mu,\delta+\nu,\nu+\beta^{**}\}}} |a^*(x,Y+g^{**}(x,Y)) \cdot \partial_x g^*(x,Y)|. \tag{79}$$

Using similar techniques we obtain

$$I^* \leqslant \frac{\nu^{-1}\Gamma(\delta)^2||a^*||_{\rho^*,\rho^{**}}||a^{**}||_{\rho^*,\rho^{**}}(\mu^{-1}||a^*||_{\rho_*,\rho^{**}} + \nu^{-1}||a^{**}||_{\rho^*,\rho^{**}})}{1-\nu^{-1}\Gamma(\delta)||a^*||_{\rho^*,\rho^{**}}}, \tag{80}$$

$$II^* \leqslant \nu^{-1}\Gamma(\delta)||a^*||_{\rho^*,\rho^{**}}||a^{**}||_{\rho^*,\rho^{**}}, \tag{81}$$

$$III^* \leqslant \mu^{-1}\Gamma(\delta)||a^*||^2_{\rho^*,\rho^{**}}. \tag{82}$$

Set $\mu = \nu$ and denote

$$||\hat{a}||_{\hat{\rho}^*,\hat{\rho}^{**}} = ||\hat{a}^*||_{\rho^*-\delta-\beta^*,\rho^{**}-\max\{\delta+\mu,\delta+\nu,\nu+\beta^{**}\}} + ||\hat{a}^{**}||_{\rho^*-\delta-\beta^*,\rho^{**}-\max\{\delta+\mu,\delta+\nu,\nu+\beta^{**}\}}, \tag{83}$$

$$||a||_{\rho^*,\rho^{**}} = ||a^*||_{\rho^*,\rho^{**}} + ||a^{**}||_{\rho^*,\rho^{**}} \tag{84}$$

and

$$T := \mu^{-1}\Gamma(\delta)||a||_{\rho^*,\rho^{**}}. \tag{85}$$

We have

$$||\hat{a}||_{\hat{\rho}^*,\hat{\rho}^{**}} \leqslant (I^{**} + II^{**}) \cdot III^{**} + I^* + II^* + III^* \leqslant \mu^{-1}\Gamma(\delta)||a||^2_{\rho^*,\rho^{**}} \cdot \left(1 + \frac{1}{2}\frac{T}{1-T}\right). \tag{86}$$

We have thus obtained quadratic estimates of the new error multiplied by terms coming from the Cauchy estimates and the estimates in the homological equation. We do not include the rest of the procedure here, as it is the same as in most papers on KAM theory. The fact that when ω is a Khintchine-Lévy number we have

$$\Gamma(\delta) = O(\delta^{-L} \log \delta^{-1}) \tag{87}$$

allows us to perform it, as the bad estimates caused by solving the homological equation and the Cauchy estimates are eventually killed by the quadraticity of the method. All we need to do is to choose δ and μ in a proper way at each step of the iterative scheme.

3 A diophantine approximation lemma

3.1 Computing $\lfloor q\omega \rfloor$ and $\lceil q\omega \rceil$

Our goal in this section is to obtain a diophantine approximation theorem, that is to get a lower bound on the quantity $|q\omega - p|$ in terms of q. First we will show that this quantity is expressible in terms of η_n's. For this we compute the p for which $|q\omega - p|$ attains its least value for a given q, that is we compute both $\lfloor q\omega \rfloor$ and $\lceil q\omega \rceil$. To do this we first introduce a coding of natural numbers into finite sequences of natural numbers with respect to a given sequence satisfying a recursive relation of the form (13). Given such a sequence $(q_n)_{n=0}^{\infty}$ we construct a finite sequence associated to a number q using the following procedure:

1. Find n such that $q_{n-1} < q \leqslant q_n$.
2. Subtract $k_{n-1}q_{n-1}$ from q_n, where k_{n-1} is the maximal number for which the result r_1 is still greater or equal than q. The number k_{n-1} satisfies $0 \leqslant k_{n-1} \leqslant a_n - 1$.
3. Subtract $k_{n-2}q_{n-2}$ from r_1, where k_{n-2} is the maximal number for which the result r_2 is still greater or equal than q. The number k_{n-2} satisfies $0 \leqslant k_{n-2} \leqslant a_{n-1}$.

...

n. Subtract k_1q_1 from r_{n-2}, where k_1 is the maximal number for which the result r_{n-1} is still greater or equal than q. The number k_1 satisfies $0 \leqslant k_1 \leqslant a_2$.

$n+1$. Subtract k_0q_0 from r_{n-1}, where k_0 is the number $r_{n-1} - q$. The number k_0 satisfies $0 \leqslant k_0 \leqslant a_1$.

The sequence $(n; k_{n-1}, \dots, k_0)$ is the desired address of q. Of course we have

$$q = q_n - (k_{n-1}q_{n-1} + k_{n-2}q_{n-2} + \dots + k_1q_1 + k_0q_0). \tag{88}$$

We define its corresponding p by replacing all the letters q in the above equality by p:

$$p = p_n - (k_{n-1}p_{n-1} + k_{n-2}p_{n-2} + \dots + k_1p_1 + k_0p_0). \tag{89}$$

Since p is, for a given q, a uniquely defined number we will sometimes refer to it as $\mathfrak{p}(q)$:

$$\mathfrak{p}(q) = p = p_n - (k_{n-1}p_{n-1} + k_{n-2}p_{n-2} + \ldots + k_1p_1 + k_0p_0). \tag{90}$$

By $\mathfrak{j}(q)$ we will mean the least index j for which k_j is nonzero and if such an index does not exist (i.e. when $q = q_n$) we define $\mathfrak{j}(q) = n+1$:

$$\mathfrak{j}(q) = \begin{cases} \min\{j : k_j \neq 0\} \text{ if } q \neq q_n \\ n+1 \text{ if } q = q_n \end{cases}. \tag{91}$$

The number $\mathfrak{j}(q)$ is again uniquely determined, which justifies the notation. We will, however, not limit ourselves to expansion of q of the form (88) given by the aforementioned procedure. In order to avoid ambiguities in the remaining part of this section we introduce the following definition.

Definition 3.1. *Assume $(q_n)_{n=0}^{\infty}$ is a sequence satisfying a recurrence relation of the form* (13) *for some sequence $(a_n)_{n=0}^{\infty}$. Identity of the form* (88) *will be called an* expansion of q *if $0 \leqslant k_m \leqslant a_{m+1}$ for $m = 0, \ldots, n-2$ and $0 \leqslant k_{n-1} \leqslant a_n - 1$.*

In principle different expansions can give rise to different functions $\mathfrak{p}$ and $\mathfrak{j}$, however this is (almost) not the case.

Lemma 3.1. *Both $\mathfrak{p}(q)$ and $(-1)^{\mathfrak{j}(q)-1}$ are independent of the expansion of q.*

We will now prove the following

Lemma 3.2. *Suppose q is a positive integer with address $(n; k_{n-1}, \ldots, k_0)$ with respect to the sequence $(q_n)_{n=0}^{\infty}$. Then $\mathfrak{p}(q) = \lfloor q\omega \rfloor$ if $\mathfrak{j}(q)$ is odd and $\mathfrak{p}(q) = \lceil q\omega \rceil$ if $\mathfrak{j}(q)$ is even.*

Before we proceed with the proof, we formulate an obvious corollary:

Corollary 3.3. *With the notations of lemma 3.2 we have the following equality of sets:*

$$\{p : |q\omega - p| < 1\} = \{\lfloor q\omega \rfloor, \lceil q\omega \rceil\} = \{\mathfrak{p}(q), \mathfrak{p}(q) + (-1)^{\mathfrak{j}(q)-1}\}. \tag{92}$$

Proof of lemma 3.2. To simplify notations we shall, for the course of the proof, replace $\mathfrak{j}(q)$ with j and $\mathfrak{p}(q)$ with p. The claim of the lemma in the case when $q = q_n$ follows from the theory of continued fractions, more specifically from a remark we made earlier that the convergents approach ω in an alternating fashion. To prove the lemma for $q \neq q_n$ it suffices to show that $0 < (-1)^{j-1}(q\omega - p) < 1$. Whenever an expansion of q appears in this proof it is the canonical expansion constructed at the beginning of this section.

$$\begin{aligned}(-1)^{j-1}(q\omega - p) &= (-1)^{j-1}[(q_n - (k_{n-1}q_{n-1} + \ldots + k_jq_j))\omega - (p_n - (k_{n-1}p_{n-1} + \ldots + k_jp_j))] = \\ &= (-1)^{j-1}[(-1)^n\eta_n - ((-1)^{n-1}k_{n-1}\eta_{n-1} + \ldots + (-1)^jk_j\eta_j)] = \\ &= k_j\eta_j - k_{j+1}\eta_{j+1} + \ldots + (-1)^{n-1-j}k_{n-1}\eta_{n-1} + (-1)^{n-1-j}\eta_n.\end{aligned} \tag{93}$$

We prove the two inequalities separately considering two cases. First, however, we introduce an auxiliary inequality:

$$\eta_s - k_{s+1}\eta_{s+1} + (k_{s+2} - 1)\eta_{s+2} \geqslant \eta_s - a_{s+2}\eta_{s+1} + (k_{s+2} - 1)\eta_{s+2} = k_{s+2}\eta_{s+2} \geqslant 0. \tag{94}$$

Case 1. The number $n-1-j$ is even.

Using equality (93) we obtain

$$\begin{aligned}(-1)^{j-1}(q\omega-p) &= k_j\eta_j - k_{j+1}\eta_{j+1} + k_{j+2}\eta_{j+2} - \ldots - k_{n-2}\eta_{n-2} + k_{n-1}\eta_{n-1} + \eta_n = \\ &= (k_j-1)\eta_j + \eta_j - k_{j+1}\eta_{j+1} + (k_{j+2}-1)\eta_{j+2} + \eta_{j+2} - k_{j+3}\eta_{j+3} + \ldots - k_{n-2}\eta_{n-2} + \\ &+ (k_{n-1}-1)\eta_{n-1} + \eta_{n-1} + \eta_n.\end{aligned} \tag{95}$$

Observe, that the first term in the above sum $(k_j-1)\eta_j$ is greater or equal than 0 (from the definition of j). Using this and our auxiliary inequality (94) applied to the splitting above we obtain

$$(-1)^{j-1}(q\omega-p) \geqslant \eta_{n-1} + \eta_n > 0. \tag{96}$$

Using similar techniques we obtain

$$\begin{aligned}(-1)^{j-1}(q\omega-p) &= k_j\eta_j - k_{j+1}\eta_{j+1} + k_{j+2}\eta_{j+2} - \ldots - k_{n-2}\eta_{n-2} + k_{n-1}\eta_{n-1} + \eta_n \leqslant \\ &\leqslant k_j\eta_j + k_{j+2}\eta_{j+2} + \ldots + k_{n-3}\eta_{n-3} + k_{n-1}\eta_{n-1}\eta_n \leqslant \\ &\leqslant a_{j+1}\eta_j + a_{j+3}\eta_{j+2} + \ldots + a_{n-2}\eta_{n-3} + (a_n-1)\eta_{n-1} + \eta_n = \\ &= \eta_{j-1} - \eta_{j+1} + \eta_{j+1} - \eta_{j+3} + \ldots + \eta_{n-4} - \eta_{n-2} + \eta_{n-2} - \eta_n - \eta_{n-1} + \eta_n = \\ &= \eta_{j-1} - \eta_{n-1} \leqslant \eta_{-1} - \eta_{n-1} = 1 - \eta_{n-1} < 1.\end{aligned} \tag{97}$$

Case 2. The number $n-1-j$ is odd.

We proceed analogously to case 1 to obtain

$$\begin{aligned}(-1)^{j-1}(q\omega-p) &= k_j\eta_j - k_{j+1}\eta_{j+1} + k_{j+2}\eta_{j+2} - \ldots + k_{n-2}\eta_{n-2} - k_{n-1}\eta_{n-1} - \eta_n = \\ &= (k_j-1)\eta_j + \eta_j - k_{j+1}\eta_{j+1} + (k_{j+2}-1)\eta_{j+2} + \eta_{j+2} - k_{j+3}\eta_{j+3} + \ldots + (k_{n-2}-1)\eta_{n-2} + \\ &+ \eta_{n-2} - k_{n-1}\eta_{n-1} - \eta_n \geqslant \eta_{n-1} > 0\end{aligned} \tag{98}$$

and

$$\begin{aligned}(-1)^{j-1}(q\omega-p) &= k_j\eta_j - k_{j+1}\eta_{j+1} + k_{j+2}\eta_{j+2} - \ldots + k_{n-2}\eta_{n-2} - k_{n-1}\eta_{n-1} - \eta_n \leqslant \\ &\leqslant k_j\eta_j + k_{j+2}\eta_{j+2} + \ldots + k_{n-4}\eta_{n-4} + k_{n-2}\eta_{n-2} \leqslant \\ &\leqslant a_{j+1}\eta_j + a_{j+3}\eta_{j+2} + \ldots + a_{n-3}\eta_{n-4} + a_{n-1}\eta_{n-2} = \\ &= \eta_{j-1} - \eta_{j+1} + \eta_{j+1} - \eta_{j+3} + \ldots + \eta_{n-5} - \eta_{n-3} + \eta_{n-3} - \eta_{n-1} = \\ &= \eta_{j-1} - \eta_{n-1} \leqslant \eta_{-1} - \eta_{n-1} = 1 - \eta_{n-1} < 1.\end{aligned} \tag{99}$$

□

3.2 A lower bound on $|q\omega-p|$ for a majority of numbers q

In this section we shall make use of lemma 3.2 to control the size of $|q\omega-p|$. The crucial observation here is that $q\omega - \mathfrak{p}(q)$ is a linear combination of η_m's with coefficients $\pm k_m$. Using this idea we shall prove the following

Theorem 3.1. *Let $\omega = [a_0; a_1, a_2, \ldots]$ be a positive irrational number and let $(q_n)_{n=0}^{\infty}$ denote the sequence of denominators of convergents to ω. Assume that $q>0$ does not belong to $ID \cup TD \cup SD \cup CSD$ and that p is an integer. Then*

$$|q\omega-p| \geqslant \frac{1}{Cq}, \tag{100}$$

where

$$C = \max\left\{2, \frac{1}{\lceil\omega\rceil - \omega}\right\}. \tag{101}$$

Before we proceed with the proof we make a remark which justifies the title of this subsection - The set $ID \cup TD \cup SD \cup CSD$ is indeed small - the slowest growth rate of the sequence (q_n) is given by the Fibonacci sequence, which is exponential. Using this it's easy to see that the excluded set has density 0 in the natural numbers.

Proof. If $|q\omega - p| > 1$ then the theorem is trivially true (for all q), thus we are left with the p's for which $|q\omega - p| < 1$. Corollary 3.3 tells us in turn that there are two such p's, namely $\mathfrak{p}(q)$ and $\mathfrak{p}(q) + (-1)^{\mathfrak{j}(q)-1}$. We shall therefore divide the proof into two cases - one dealing with $|q\omega - \mathfrak{p}(q)|$ and the other with $|q\omega - \big(\mathfrak{p}(q) + (-1)^{\mathfrak{j}(q)-1}\big)|$ (both of which will be further divided into subcases). Again, in order to simplify the notations, we replace $\mathfrak{p}(q)$ and $\mathfrak{j}(q)$ by p and j for the course of the proof, since this causes no ambiguities.

The main idea now is to show that in both cases we can get a lower bound by η_{n-2} multiplied by some constant (n being the first digit in the address of q, i.e. such that $q_{n-1} < q \leqslant q_n$) provided that q does not belong to the four excluded sets. This is enough, since

$$\eta_{n-2} = |q_{n-2}\omega - p_{n-2}| > \frac{1}{q_{n-1} + q_{n-2}} > \frac{1}{2q}. \tag{102}$$

Observe that in lemma 3.2 we actually showed that

$$|q\omega - p| \geqslant \eta_{n-1}. \tag{103}$$

It is the transition from η_{n-1} to η_{n-2} in the above inequality that forces us to reject some numbers q from our considerations.

In this proof we do not stick to one particular expansion of q, instead we frequently switch between different expansions. Observe however, that in view of lemma 3.1 this is justified, since - as will turn out in the course of the proof - all essential quantities appearing along the way depend only on p and $(-1)^{j-1}$ and are thus do not depend on the choice of expansion.

Case 1. *Estimating* $|q\omega - p|$.

Case 1.1. *The number* $n - 1 - j$ *is even.*

$$\begin{aligned}
|q\omega - p| &= (-1)^{j-1}(q\omega - p) = \\
&= k_j\eta_j - k_{j+1}\eta_{j+1} + k_{j+2}\eta_{j+2} - \ldots - k_{n-2}\eta_{n-2} + k_{n-1}\eta_{n-1} + \eta_n = \\
&= (k_j - 1)\eta_j + \eta_j - k_{j+1}\eta_{j+1} + (k_{j+2} - 1)\eta_{j+2} + \eta_{j+2} - k_{j+3}\eta_{j+3} + \\
&+ \ldots - k_{n-2}\eta_{n-2} + (k_{n-1} - 1)\eta_{n-1} + \eta_{n-1} + \eta_n = \\
&= (k_j - 1)\eta_j + (a_{j+2} - k_{j+1})\eta_{j+1} + k_{j+2}\eta_{j+2} + (a_{j+4} - k_{j+3})\eta_{j+3} + k_{j+4}\eta_{j+4} + \\
&+ \ldots + (a_{n-1} - k_{n-2})\eta_{n-2} + k_{n-1}\eta_{n-1} + \eta_{n-1} + \eta_n.
\end{aligned} \tag{104}$$

In the above we have used the fact that

$$\eta_s - k_{s+1}\eta_{s+1} + (k_{s+2} - 1)\eta_{s+2} = (a_{s+2} - k_{s+1})\eta_{s+1} + k_{s+2}\eta_{s+2}. \tag{105}$$

Observe that the last expression in (104) can be not greater than η_{n-2} only if $k_j = 1$, $k_{j+1} = a_{j+2}$, $k_{j+2} = 0$, $k_{j+3} = a_{j+4}$, $\ldots$, $k_{n-2} = a_{n-1}$ and k_{n-1} is arbitrary. Otherwise, there will always emerge an η_s with $s \leqslant n - 2$

in this sum. Numbers q with this type of address are equal to

$$\begin{aligned} q &= q_n - k_{n-1}q_{n-1} - a_{n-1}q_{n-2} - a_{n-3}q_{n-4} - \ldots - a_{j+2}q_{j+1} - q_j = \\ &= q_n - k_{n-1}q_{n-1} - (q_{n-1} - q_{n-3} + q_{n-3} - q_{n-5} + \ldots + q_{j+2} - q_j + q_j) = \\ &= q_n - (k_{n-1}+1)q_{n-1} = q_{n-2} + (a_n - k_{n-1} - 1)q_{n-1}, \end{aligned} \tag{106}$$

that is are of the form $q_{n-2} + aq_{n-1}$ for some $a = 0, \ldots, a_n - 1$, thus they are among the excluded q's - either initial denominators, true denominators or semidenominators.

Case 1.2. *The number $n-1-j$ is odd.*

Just like in case 1.1 we obtain that if we want to have $|q\omega - p| \geqslant \eta_{n-2}$ we need to exclude addresses of the form

$$(n; k_{n-1}, 0, a_{n-2}, 0, a_{n-4}, 0, \ldots, 0, a_{j+4}, 0, a_{j+2}, 1, 0, 0, 0, \ldots, 0, 0) \tag{107}$$

with arbitrary k_{n-1}.

$$\begin{aligned} |q\omega - p| &= (-1)^{j-1}(q\omega - p) = \\ &= k_j\eta_j - k_{j+1}\eta_{j+1} + k_{j+2}\eta_{j+2} - \ldots + k_{n-2}\eta_{n-2} - k_{n-1}\eta_{n-1} - \eta_n = \\ &= (k_j - 1)\eta_j + \eta_j - k_{j+1}\eta_{j+1} + (k_{j+2} - 1)\eta_{j+2} + \eta_{j+2} - k_{j+3}\eta_{j+3} + \\ &+ \ldots + k_{n-2}\eta_{n-2} - k_{n-1}\eta_{n-1} + \eta_n = \\ &= (k_j - 1)\eta_j + (a_{j+2} - k_{j+1})\eta_{j+1} + k_{j+2}\eta_{j+2} + (a_{j+4} - k_{j+3})\eta_{j+3} + k_{j+4}\eta_{j+4} + \\ &+ \ldots + k_{n-2}\eta_{n-2} + (a_n - k_{n-1})\eta_{n-1}. \end{aligned} \tag{108}$$

Numbers q with this type of address are equal to

$$\begin{aligned} q &= q_n - k_{n-1}q_{n-1} - a_{n-2}q_{n-3} - a_{n-4}q_{n-5} - \ldots - a_{j+2}q_{j+1} - q_j = \\ &= q_n - k_{n-1}q_{n-1} - (q_{n-2} - q_{n-4} + q_{n-4} - q_{n-6} + \ldots + q_{j+2} - q_j + q_j) = \\ &= q_n - k_{n-1}q_{n-1} - q_{n-2} = (a_n - k_{n-1})q_{n-1}, \end{aligned} \tag{109}$$

that is of the form aq_{n-1} for some $a = 1, \ldots, a_n$, thus they are among the excluded q's - either initial denominators, true denominators or cosemidenominators.

Case 2. *Estimating $\left|q\omega - \left(p + (-1)^{j-1}\right)\right|$.*

First note that since $|q\omega - p| < 1$ we have

$$\begin{aligned} |q\omega - (p + (-1)^{j-1})| &= |q\omega - p + (-1)^j| = |(-1)^{j-1}|q\omega - p| + (-1)^j| = \\ &= ||q\omega - p| - 1| = 1 - |q\omega - p| = 1 - (-1)^{j-1}(q\omega - p). \end{aligned} \tag{110}$$

Case 2.1. *The number $n-1-j$ is even.*

We will use the following representation of 1 as a sum (its proof follows from identity (15)):

$$1 = 1 - \eta_{j-1} + a_{j+1}\eta_j + a_{j+3}\eta_{j+2} + \ldots + a_{n-2}\eta_{n-3} + a_n\eta_{n-1} + \eta_n. \tag{111}$$

Now, using the above, we obtain

$$\begin{aligned} 1 - (-1)^{j-1}(q\omega - p) &= 1 - k_j\eta_j + k_{j+1}\eta_{j+1} - k_{j+2}\eta_{j+2} + \ldots + k_{n-2}\eta_{n-2} - k_{n-1}\eta_{n-1} - \eta_n = \\ &= 1 - \eta_{j-1} + (a_{j+1} - k_j)\eta_j + k_{j+1}\eta_{j+1} + (a_{j+3} - k_{j+2})\eta_{j+2} + \ldots + \\ &+ k_{n-2}\eta_{n-2} + (a_n - (k_{n-1}+1))\eta_{n-1} + \eta_{n-1}. \end{aligned} \tag{112}$$

Note that because of the above the minimal value of $1 - (-1)^{j-1}(q\omega - p)$ is equal to $1 - \eta_{j-1} + \eta_{n-1}$ and is attained only if $k_{n-1} = a_n - 1$, $k_{n-2} = 0$, $k_{n-3} = a_{n-2}$, $k_{n-4} = 0$, $\ldots$, $k_{j+2} = a_{j+3}$, $k_{j+1} = 0$ and $k_j = a_{j+1}$.

Case 2.1.1. *The number k_{n-1} is arbitrary, while $k_{n-2} = 0$, $k_{n-3} = a_{n-2}$, $k_{n-4} = 0$, $\ldots$, $k_{j+2} = a_{j+3}$, $k_{j+1} = 0$ and $k_j = a_{j+1}$.*

We will try to estimate $1 - \eta_{j-1} + (a_n - (k_{n-1} + 1))\eta_{n-1}$, or actually even $1 - \eta_{j-1}$ from below by η_{n-2}.

Case 2.1.1.1. $j \geqslant 2$.

Observe that this implies $n \geqslant 2$.

Case 2.1.1.1.1. $j - 1$ *is odd.*

$$\begin{aligned} 1 - \eta_{j-1} &= \eta_{-1} - \eta_1 + \eta_1 - \eta_3 + \ldots + \eta_{j-3} - \eta_{j-1} = \\ &= a_1\eta_0 + a_3\eta_2 + \ldots + a_{j-1}\eta_{j-2} \geqslant a_1\eta_0 \geqslant \eta_0 \geqslant \eta_{n-2}. \end{aligned} \tag{113}$$

Case 2.1.1.1.2. $j - 1$ *is even.*

$$\begin{aligned} 1 - \eta_{j-1} &= 1 - \eta_0 + \eta_0 - \eta_2 + \eta_2 - \eta_4 + \ldots + \eta_{j-3} - \eta_{j-1} = \\ &= 1 - \eta_0 + a_2\eta_1 + a_4\eta_3 + \ldots + a_{j-1}\eta_{j-2} \geqslant \\ &\geqslant 1 - \eta_0 + a_2\eta_1 = 1 - \eta_2 > 1 - \frac{1}{q_3} \geqslant \frac{2}{3} > \frac{1}{2q}. \end{aligned} \tag{114}$$

Case 2.1.1.2. $j = 1$.

$$1 - \eta_0 = \lceil\omega\rceil - \omega \geqslant \frac{1}{C_{2112}q}, \tag{115}$$

where $C_{2112} = \frac{1}{\lceil\omega\rceil - \omega}$.

Case 2.1.1.3. $j = 0$.

All the numbers q of this case have address of the form

$$(n; k_{n-1}, 0, a_{n-2}, 0, a_{n-4}, 0, \ldots, a_5, 0, a_3, 0, a_1) \tag{116}$$

with n odd and k_{n-1} arbitrary ($0 \leqslant k_{n-1} \leqslant a_n - 1$ as always). These numbers are equal to

$$\begin{aligned} q &= q_n - (k_{n-1}q_{n-1} + a_{n-2}q_{n-3} + a_{n-4}q_{n-5} + \ldots + a_5q_4 + a_3q_2 + a_1q_0) = \\ &= q_n - (k_{n-1}q_{n-1} + q_{n-2} - q_{n-4} + q_{n-4} - q_{n-6} + \ldots + q_5 - q_3 + q_3 - q_1 + q_1 - q_{-1}) = \\ &= q_n - q_{n-2} - k_{n-1}q_{n-1} = (a_n - k_{n-1})q_{n-1}. \end{aligned} \tag{117}$$

We see that, as in case 1.2, they are among the numbers, which have been excluded in the assumptions (among initial denominators, true denominators or cosemidenominators).

Case 2.1.2. *One of the equalities of case 2.1.1 does not hold.*

In this case in expression (112) there will always emerge an η_s with $s \leqslant n - 2$. Since $1 - \eta_{j-1} \geqslant 0$ we can bound everything from below by η_{n-2}, which is what we want in view of inequality (102).

Case 2.2. *The number $n-1-j$ is odd.*

Our reasoning will be very similar to the one in case 2.1. We first use an analogue of (111):

$$1 = 1 - \eta_{j-1} + a_{j+1}\eta_j + a_{j+3}\eta_{j+2} + \ldots + a_{n-3}\eta_{n-4} + a_{n-1}\eta_{n-2} + \eta_{n-1}. \tag{118}$$

Now, as before, we obtain

$$\begin{aligned}1-(-1)^{j-1}(q\omega-p) &= 1 - k_j\eta_j + k_{j+1}\eta_{j+1} - k_{j+2}\eta_{j+2} + \ldots + k_{n-3}\eta_{n-3} - k_{n-2}\eta_{n-2} + k_{n-1}\eta_{n-1} + \eta_n = \\ &= 1 - \eta_{j-1} + (a_{j+1}-k_j)\eta_j + k_{j+1}\eta_{j+1} + (a_{j+3}-k_{j+2})\eta_{j+2} + \ldots + \\ &+ k_{n-3}\eta_{n-3} + (a_{n-1}-k_{n-2})\eta_{n-2} + k_{n-1}\eta_{n-1} + \eta_n.\end{aligned} \tag{119}$$

The minimal value of expression (119) is equal to $1-\eta_{j-1}+\eta_n$ and is attained when the address of q is equal to

$$(n; 0, a_{n-1}, 0, a_{n-3}, 0, \ldots, 0, a_{j+3}, 0, a_{j+1}, 0, 0, \ldots, 0, 0), \tag{120}$$

where a_{j+1} stands in the j-th place from the right (as usual the last position in the sequence is the zeroth place from the right).

Case 2.2.1. *The number k_{n-1} is arbitrary, while $k_{n-2} = a_{n-1}$, $k_{n-3} = 0$, $k_{n-4} = a_{n-3}$, $k_{n-5} = 0$, $\ldots$, $k_{j+2} = a_{j+3}$, $k_{j+1} = 0$, $k_j = a_{j+1}$.*

We will try to estimate $1-\eta_{j-1}+\eta_n$ from below by η_{n-2}.

Case 2.2.1.1. $j \geqslant 2$.

Reasoning in this case is identical to the one in case 2.1.1.1.

Case 2.2.1.2. $j = 1$.

Reasoning in this case is identical to the one in case 2.1.1.2.

Case 2.2.1.3. $j = 0$.

We will once more determine all numbers q which satisfy the conditions of this case. These numbers have addresses of the form

$$(n; k_{n-1}, a_{n-1}, 0, a_{n-3}, 0, a_{n-5}, \ldots, a_5, 0, a_3, 0, a_1) \tag{121}$$

with n even and k_{n-1} arbitrary. These numbers are equal to

$$\begin{aligned}q &= q_n - (k_{n-1}q_{n-1} + a_{n-1}q_{n-2} + a_{n-3}q_{n-4} + \ldots + a_5q_4 + a_3q_2 + a_1q_0) = \\ &= q_n - (k_{n-1}q_{n-1} + q_{n-1} - q_{n-3} + q_{n-3} - q_{n-5} + \ldots + q_5 - q_3 + q_3 - q_1 + q_1 - q_{-1}) = \\ &= q_n - q_{n-1} - k_{n-1}q_{n-1} = q_{n-2} + (a_n - (k_{n-1}+1))q_{n-1}.\end{aligned} \tag{122}$$

We see that all such q's are among the excluded ones (among initial denominators, true denominators or semidenominators).

Case 2.2.2. *One of the equalities of case 2.2.1 does not hold.*

As in case 2.1.2 in expression (119) there will emerge an η_s with $s \leqslant n-2$ and this fact combined with inequality $1-\eta_{j-1} \geqslant 0$ will result in a lower bound by η_{n-2}.

□

We remark that apart from the result in the statement of theorem 3.2 the methodology used in the proof can give a detailed insight into the dependence of actual size of $|q\omega - p|$ on q. There is room for improvement, one could refine the result (which would mean an even more nested structure of cases in the proof) to obtain far more precise estimates. This could be done, for instance, by not focusing on bounding everything in the proof from below by η_{n-2} but by η_{n-3}. However we do not pursue this here, as theorem 3.2 will be sufficiently good for our applications.

It turns out that some simple calculations can provide us with a slight generalization of theorem 3.2. To formulate it we first introduce some notations:

$$TP = \{(q,p) \in \mathbb{Z}^2_* : q = q_n \text{ and } p = p_n \text{ for some } n \geqslant 2\}, \tag{123}$$

$$SP = \{(q,p) \in \mathbb{Z}^2_* : q = aq_{n-1} + q_{n-2} \text{ and } p = ap_{n-1} + p_{n-2} \text{ for some } n \geqslant 2 \text{ and } a \in \{1, \ldots, a_n - 1\}\}, \tag{124}$$

$$CSP = \{(q,p) \in \mathbb{Z}^2_* : q = aq_{n-1} \text{ and } p = ap_{n-1} \text{ for some } n \geqslant 2 \text{ and } a \in \{2, \ldots, a_n\}\}, \tag{125}$$

$$IP = \{(q,p) \in \mathbb{Z}^2_* : 1 \leqslant q \leqslant a_1 \text{ and } p = 1 + qa_0\}, \tag{126}$$

$$IP' = \{(q,p) \in \mathbb{Z}^2_* : 1 \leqslant q \leqslant a_1 \text{ and } p = qa_0\}, \tag{127}$$

$$\text{Bad} = TP \cup -TP \cup SP \cup -SP \cup CSP \cup -CSP \cup IP \cup -IP \cup IP' \cup -IP', \tag{128}$$

$$\text{Good} = \mathbb{Z}^2_* \setminus \text{Bad}. \tag{129}$$

Theorem 3.2. *Suppose ω is irrational and $(q,p) \in$ Good. Then*

$$|q\omega - p| \geqslant \frac{1}{Cq}, \tag{130}$$

where C is the constant from theorem 3.1.

4 Bounds in the homological equation

4.1 Preparatory estimations

In this section we consider the homological equation suitable for the problem of rectification of a perturbation of a constant torus flow and we give bounds on its solution in terms of the initial data. More specifically let $\delta > 0$ be a number smaller than ρ. The goal of this section is to find estimates on $||h||_{\rho-\delta}$ in terms of δ and $||a||_\rho$. In order to do this we first need a lemma about the decay rate of Fourier coefficients of the function a:

Lemma 4.1. *Let $f \in \mathcal{P}_\rho$ and let f_l denote its Fourier expansion terms. Then*

$$|f_l| \leqslant ||f||_\rho \cdot e^{-|l|\rho}.$$

Proof. We shall follow the exposition given in [3]. The Fourier series terms are given by the equality

$$f_{l_1,l_2} = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} f(x,y) e^{-i(l_1 x + l_2 y)} dx\, dy,$$

so we need to estimate the absolute value of this integral. To do this we shall use the fact that in the case of analytic functions we have the freedom to change the path of integration. Instead of integrating over the segment between 0 and 2π we shall integrate over the sum of three segments: $[0, \pm i\rho]$, $[\pm i\rho, 2\pi \pm i\rho]$ and $[2\pi \pm i\rho, 2\pi]$, where the signs will be chosen in accordance with the signs of l_1 and l_2. The integrals over the first and third interval will cancel out because of the periodicity of the integrand in real directions.

$$
\begin{aligned}
|f_{l_1,l_2}| &= \frac{1}{(2\pi)^2}\left|\int_{y=0}^{y=2\pi}\int_{x=0}^{x=2\pi} f(x,y)e^{-i(l_1x+l_2y)}dx\,dy\right| = \\
&= \frac{1}{(2\pi)^2}\left|\int_{y=0}^{y=2\pi} e^{-il_2y}\cdot\left[\left(\int_{x=0}^{x=-i\rho\cdot\operatorname{sgn} l_1}+\int_{x=-i\rho\cdot\operatorname{sgn} l_1}^{x=2\pi-i\rho\cdot\operatorname{sgn} l_1}+\int_{x=2\pi-i\rho\cdot\operatorname{sgn} l_1}^{x=2\pi}\right)f(x,y)e^{-l_1x}dx\right]dy\right| = \\
&= \frac{1}{(2\pi)^2}\left|\int_{y=0}^{y=2\pi} e^{-il_2y}\cdot\left[\left(\int_{x=0}^{x=-i\rho\cdot\operatorname{sgn} l_1}+\int_{x=-i\rho\cdot\operatorname{sgn} l_1}^{x=2\pi-i\rho\cdot\operatorname{sgn} l_1}-\int_{x=2\pi}^{x=2\pi-i\rho\cdot\operatorname{sgn} l_1}\right)f(x,y)e^{-l_1x}dx\right]dy\right| = \\
&= \frac{1}{(2\pi)^2}\left|\int_{y=0}^{y=2\pi} e^{-il_2y}\cdot\left[\int_{x=-i\rho\cdot\operatorname{sgn} l_1}^{x=2\pi-i\rho\cdot\operatorname{sgn} l_1}f(x,y)e^{-l_1x}dx\right]dy\right| = \\
&= \frac{1}{(2\pi)^2}\left|\int_{y=-i\rho\cdot\operatorname{sgn} l_2}^{y=2\pi-i\rho\cdot\operatorname{sgn} l_2} e^{-il_2y}\cdot\left[\int_{x=-i\rho\cdot\operatorname{sgn} l_1}^{x=2\pi-i\rho\cdot\operatorname{sgn} l_1}f(x,y)e^{-l_1x}dx\right]dy\right| \leqslant \qquad (131)\\
&\leqslant \frac{1}{(2\pi)^2}\cdot 2\pi\cdot\sup_{y\in[-i\rho\cdot\operatorname{sgn} l_2,2\pi-i\rho\cdot\operatorname{sgn} l_2]}\left|e^{-il_2y}\left[\int_{-i\rho\cdot\operatorname{sgn} l_1}^{2\pi-i\rho\cdot\operatorname{sgn} l_1}f(x,y)e^{-il_1x}dx\right]\right| \leqslant \\
&\leqslant \frac{1}{2\pi}e^{-|l_2|\rho}\cdot\sup_{y\in\Pi(\rho)}\left|\int_{-i\rho\cdot\operatorname{sgn} l_1}^{2\pi-i\rho\cdot\operatorname{sgn} l_1}f(x,y)e^{-il_1x}dx\right| \leqslant \\
&\leqslant \frac{1}{2\pi}e^{-|l_2|\rho}\cdot\sup_{y\in\Pi(\rho)}\left[2\pi\cdot\sup_{x\in[-i\rho\cdot\operatorname{sgn} l_1,2\pi-i\rho\cdot\operatorname{sgn} l_1]}\left|f(x,y)e^{-il_1x}\right|\right] \leqslant \\
&\leqslant e^{-|l_2|\rho}\cdot e^{-|l_1|\rho}\sup_{y\in\Pi(\rho)}\sup_{x\in\Pi(\rho)}|f(x,y)| = ||f||_\rho\cdot e^{-|l|\rho}.
\end{aligned}
$$

□

We will now make an attempt to estimate $||h||_{\rho-\delta}$:

$$
\begin{aligned}
||h||_{\rho-\delta} = \sup_{x,y\in\Pi(\rho)}|h(x,y)| &= \sup_{x,y\in\Pi(\rho-\delta)}\left|\sum_{l\in\mathbb{Z}^2_*}\frac{a_l}{i(l_1+l_2\omega)}e^{i(l_1x+l_2y)}\right| \leqslant \\
&\leqslant \sup_{x,y\in\Pi(\rho-\delta)}\left|\sum_{l:(l_2,-l_1)\in\mathrm{Good}}\frac{a_l}{i(l_1+l_2\omega)}e^{i(l_1x+l_2y)}\right| + \sum_{l:(l_2,-l_1)\in\mathrm{Bad}}\left(\frac{|a_l|}{|l_1+l_2\omega|}\cdot\sup_{x,y\in\Pi(\rho-\delta)}\left|e^{i(l_1x+l_2y)}\right|\right) = \\
&= \sup_{x,y\in\Pi(\rho-\delta)}\left|\sum_{l:(l_2,-l_1)\in\mathrm{Good}}\frac{a_l}{i(l_1+l_2\omega)}e^{i(l_1x+l_2y)}\right| + \sum_{l:(l_2,-l_1)\in\mathrm{Bad}}\frac{|a_l|}{|l_1+l_2\omega|}e^{|l|(\rho-\delta)} \overset{(\star)}{\leqslant} \\
&\leqslant \sup_{x,y\in\Pi(\rho-\delta)}\left|\sum_{l:(l_2,-l_1)\in\mathrm{Good}}\frac{a_l}{i(l_1+l_2\omega)}e^{i(l_1x+l_2y)}\right| + ||a||_\rho\sum_{l:(l_2,-l_1)\in\mathrm{Bad}}\frac{e^{-|l|\delta}}{|l_1+l_2\omega|}.
\end{aligned}
\qquad (132)
$$

In inequality $(\star)$ we used lemma 4.1. We performed the splitting above for a reason - the first summand in the above sum can be now bounded from above using the following theorem of Rüssmann, which we state in a form which suits best our needs:

Theorem 4.1 (Rüssmann, [25])**.** *Suppose $a \in \mathcal{P}_\rho$ for some $\rho > 0$ and that the average of a is zero. Assume also that ω is such that the inequalities*

$$|q\omega - p| \geqslant \frac{1}{Cq} \tag{133}$$

hold for all $(q,p) \in \mathbb{Z}^2_*$. Then the differential equation (HOM) *has a unique zero-mean solution h which belongs to $\mathcal{P}_{\rho-\delta}$ for any $0 < \delta < \rho$. Moreover, we have*

$$||h||_{\rho-\delta} \leqslant \frac{9\pi}{2} C\delta^{-1} ||a||_\rho. \tag{134}$$

The method of proof of theorem 4.1 allows us, however, to draw a conclusion useful for us. Namely if we assume that the diophantine condition is satisfied only for $(q,p) \in$ Good and consider the projection of h onto the subspace spanned by the Good elements of the Fourier basis, then the theorem still holds. This way we are left with estimating

$$\sum_{(q,p)\in \mathrm{Bad}} \frac{e^{-(|q|+|p|)\delta}}{|q\omega - p|}. \tag{135}$$

4.2 Estimates with a Brjuno-like function

We will now provide bounds for the sum over Bad and for this we introduce Brjuno-like functions. We will introduce them after performing some motivating computations. To give bounds on the sum over Bad we first observe that for $J \subset \mathbb{Z}^2_*$ the sums over J and $-J$ are equal, since the summands are symmetric. This way

$$\sum_{\mathrm{Bad}} = 2 \cdot \left(\sum_{IP} + \sum_{IP'} + \sum_{TP} + \sum_{SP} + \sum_{CSP} \right) \tag{136}$$

We will give bounds on each of the above sums.

1. Sum over TP.

$$\sum_{(q,p)\in TP} \frac{e^{-(|q|+|p|)\delta}}{|q\omega - p|} = \sum_{n\geqslant 2} \frac{e^{-(q_n+p_n)\delta}}{\eta_n} \tag{137}$$

From lemma 3.2 we know that p_n is either equal to $\lfloor q_n\omega \rfloor$ or $\lceil q\omega \rceil$, therefore it is certainly bounded from below by $q_n\omega - 1$. Thus

$$\sum_{n\geqslant 2} \frac{e^{-(q_n+p_n)\delta}}{\eta_n} \leqslant \sum_{n\geqslant 2} \frac{e^{-(q_n+q_n\omega-1)\delta}}{\eta_n} = e^\delta \sum_{n\geqslant 2} \frac{e^{-q_n(1+\omega)\delta}}{\eta_n}. \tag{138}$$

We define $\mathsf{brj}^1_\omega(\delta)$ to be the last sum:

$$\mathsf{brj}^1_\omega(\delta) = e^\delta \sum_{n\geqslant 2} \frac{e^{-q_n(1+\omega)\delta}}{\eta_n}. \tag{139}$$

2. Sum over SP.

$$\sum_{(q,p)\in SP} \frac{e^{-(|q|+|p|)\delta}}{|q\omega - p|} = \sum_{n\geqslant 2} \sum_{a=1}^{a_n-1} \frac{e^{-(aq_{n-1}+q_{n-2}+ap_{n-1}+p_{n-2})\delta}}{|a(q_{n-1}\omega - p_{n-1}) + (q_{n-2}\omega - p_{n-2})|} \tag{140}$$

Again, $ap_{n-1}+p_{n-2}$ is either equal to the floor or ceiling of $(aq_{n-1}+q_{n-2})\omega$, thus it is bounded from below by $(aq_{n-1}+q_{n-2})\omega-1$. Thus

$$\sum_{n\geqslant 2}\sum_{a=1}^{a_n-1}\frac{e^{-(aq_{n-1}+q_{n-2}+ap_{n-1}+p_{n-2})\delta}}{|a(q_{n-1}\omega-p_{n-1})+(q_{n-2}\omega-p_{n-2})|}\leqslant e^{\delta}\sum_{n\geqslant 2}\sum_{a=1}^{a_n-1}\frac{e^{-(aq_{n-1}+q_{n-2})(1+\omega)\delta}}{\eta_{n-2}-a\eta_{n-1}}= \tag{141}$$

$$=e^{\delta}\sum_{n\geqslant 2}\sum_{a=1}^{a_n-1}\frac{e^{-(aq_{n-1}+q_{n-2})(1+\omega)\delta}}{\eta_n+a_n\eta_{n-1}-a\eta_{n-1}}\overset{a\mapsto a_n-a}{=} \tag{142}$$

$$=e^{\delta}\sum_{n\geqslant 2}\sum_{a=1}^{a_n-1}\frac{e^{-(a_nq_{n-1}-aq_{n-1}+q_{n-2})(1+\omega)\delta}}{\eta_n+a\eta_{n-1}}\leqslant \tag{143}$$

$$\leqslant e^{\delta}\sum_{n\geqslant 2}\frac{e^{-(a_nq_{n-1}+q_{n-2})(1+\omega)\delta}}{\eta_{n-1}}\cdot\sum_{a=1}^{a_n-1}\frac{1}{a}e^{aq_{n-1}(1+\omega)\delta}\leqslant \tag{144}$$

$$\leqslant e^{\delta}\sum_{n\geqslant 2}\frac{e^{-(q_n-(a_n-1)q_{n-1})(1+\omega)\delta}}{\eta_{n-1}}\cdot H_{a_n-1}= \tag{145}$$

$$=e^{\delta}\sum_{n\geqslant 1}\frac{e^{-(q_{n-1}+q_n)(1+\omega)\delta}}{\eta_n}H_{a_{n+1}-1}, \tag{146}$$

where H_m denotes the m-th harmonic number:

$$H_m=\begin{cases}0 \text{ if } m=0\\ \sum_{j=1}^{m}j^{-1} \text{ if } m>0\end{cases}. \tag{147}$$

We define $\mathsf{brj}^2_\omega(\delta)$ to be the last sum:

$$\mathsf{brj}^2_\omega(\delta)=e^{\delta}\sum_{n\geqslant 1}\frac{e^{-(q_{n-1}+q_n)(1+\omega)\delta}}{\eta_n}H_{a_{n+1}-1}. \tag{148}$$

3. Sum over CSP. Analogously to the case of TP and SP we obtain

$$\sum_{(q,p)\in CSP}\frac{e^{-(|q|+|p|)\delta}}{|q\omega-p|}=\sum_{n\geqslant 2}\sum_{a=2}^{a_n}\frac{e^{-(aq_{n-1}+ap_{n-1})\delta}}{a|q_{n-1}\omega-p_{n-1}|}\leqslant e^{\delta}\sum_{n\geqslant 2}\sum_{a=2}^{a_n}\frac{e^{-aq_{n-1}(1+\omega)\delta}}{a\eta_{n-1}}\leqslant e^{\delta}\sum_{n\geqslant 1}\frac{e^{-2q_n(1+\omega)\delta}}{\eta_n}(H_{a_n}-1). \tag{149}$$

We define $\mathsf{brj}^3_\omega(\delta)$ to be the last sum:

$$\mathsf{brj}^3_\omega(\delta)=e^{\delta}\sum_{n\geqslant 1}\frac{e^{-2q_n(1+\omega)\delta}}{\eta_n}(H_{a_{n+1}}-1). \tag{150}$$

Definition 4.1. *We call a number $\omega\in\mathbb{R}\setminus\mathbb{Q}$ super-Brjuno if for all $\delta>0$ the series $\mathsf{brj}^1_\omega(\delta),\mathsf{brj}^2_\omega(\delta)$ and $\mathsf{brj}^3_\omega(\delta)$ are convergent.*

Definition 4.2. *We define the super-Brjuno function to be*

$$\mathsf{brj}_\omega(\delta)=\mathsf{brj}^1_\omega(\delta)+\mathsf{brj}^2_\omega(\delta)+\mathsf{brj}^3_\omega(\delta). \tag{151}$$

Collectively we have thus proved above the following

Theorem 4.2. *Suppose ω is a super-Brjuno number. Then the solution h of the homological equation* (HOM) *satisfies*

$$||h||_{\rho-\delta}||a||_{\rho}^{-1} \leqslant \frac{9\pi}{2}C\delta^{-1} + 2\mathsf{brj}_{\omega}(\delta) + 2\sum_{q=1}^{a_1}\frac{e^{-(q+qa_0+1)\delta}}{|q\omega - qa_0 - 1|} + 2\sum_{q=1}^{a_1}\frac{e^{-q(1+a_0)\delta}}{q|\omega - a_0|} \tag{152}$$

4.3 Estimates under the Khintchine - Lévy condition

We now restrict our attention to Khintchine-Lévy numbers - from now until the end of this section we assume that $\omega \in KL(T_-, T_+, N)$ for some $T_-, T_+ > 0$ and $N \in \mathbb{N}$. Our goal is to estimate $\mathsf{brj}_{\omega}(\delta)$ (and thus in view of theorem 4.2 on $||h||_{\rho-\delta}$). For convenience we first introduce the "finite" and "tail" parts of the super-Brjuno functions and a technical lemma.

Definition 4.3. *We define*

$$\mathsf{brj}_{\omega}^{1(N)}(\delta) = e^{\delta}\sum_{n=2}^{N-1}\frac{e^{-q_n(1+\omega)\delta}}{\eta_n}, \tag{153}$$

$$\mathsf{brj}_{\omega}^{2(N)}(\delta) = e^{\delta}\sum_{n=1}^{N-1}\frac{e^{-(q_{n-1}+q_n)(1+\omega)\delta}}{\eta_n}H_{a_{n+1}-1}, \tag{154}$$

$$\mathsf{brj}_{\omega}^{3(N)}(\delta) = e^{\delta}\sum_{n=1}^{N-1}\frac{e^{-2q_n(1+\omega)\delta}}{\eta_n}(H_{a_{n+1}} - 1) \tag{155}$$

$$\mathsf{brj}_{\omega}^{1[N]}(\delta) = e^{\delta}\sum_{n=N}^{\infty}\frac{e^{-q_n(1+\omega)\delta}}{\eta_n}, \tag{156}$$

$$\mathsf{brj}_{\omega}^{2[N]}(\delta) = e^{\delta}\sum_{n=N}^{\infty}\frac{e^{-(q_{n-1}+q_n)(1+\omega)\delta}}{\eta_n}H_{a_{n+1}-1}, \tag{157}$$

$$\mathsf{brj}_{\omega}^{3[N]}(\delta) = e^{\delta}\sum_{n=N}^{\infty}\frac{e^{-2q_n(1+\omega)\delta}}{\eta_n}(H_{a_{n+1}} - 1). \tag{158}$$

and

$$\mathsf{brj}_{\omega}^{(N)}(\delta) = \mathsf{brj}_{\omega}^{1(N)}(\delta) + \mathsf{brj}_{\omega}^{2(N)}(\delta) + \mathsf{brj}_{\omega}^{3(N)}(\delta), \tag{159}$$

$$\mathsf{brj}_{\omega}^{[N]}(\delta) = \mathsf{brj}_{\omega}^{1[N]}(\delta) + \mathsf{brj}_{\omega}^{2[N]}(\delta) + \mathsf{brj}_{\omega}^{3[N]}(\delta). \tag{160}$$

Lemma 4.2. *Let m be a natural number. The following inequalities hold:*

$$a_{m+1} < \frac{q_{m+1}}{q_m}, \tag{161}$$

$$H_m - 1 \leqslant \log m, \tag{162}$$

$$H_{m-1} \leqslant \frac{1}{\log 2}\log m. \tag{163}$$

Proof. As for the first inequality we have

$$q_{m+1} = a_{m+1}q_m + q_{m-1} > a_{m+1}q_m. \tag{164}$$

As for the second one

$$\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{m}<\int_1^2\frac{dx}{x}+\int_2^3\frac{dx}{x}+\ldots+\int_{m-1}^m\frac{dx}{x}=\log m. \tag{165}$$

As for the third observe that it equality holds for $m=1,2$ and for $m\geqslant 3$ we have

$$H_{m-1}=1+H_{m-1}-1\leqslant 1+\log(m-1)\leqslant\frac{1}{\log 2}\log m. \tag{166}$$

□

Now since the Khintchine-Lévy condition deals with q_n for $n\geqslant N$ we will give bounds on the "tail" parts of the super-Brjuno functions. For simplicity denote

$$\Delta=(1+\omega)\delta. \tag{167}$$

In what follows by Γ we will mean the Euler Γ function and by Γ' its derivative.

1. Bounds on $\mathsf{brj}_\omega^{1[N]}(\delta)$*.*

$$e^{-\delta}\mathsf{brj}_\omega^{1[N]}(\delta)=\sum_{n=N}^{\infty}\frac{e^{-q_n\Delta}}{\eta_n}\leqslant\sum_{n=N}^{\infty}q_{n+1}e^{-q_n\Delta}\leqslant\sum_{n=N}e^{(\ell+T_+)(n+1)}\exp\left(-\frac{1}{2}e^{(\ell-T_-)n}\Delta\right)= \tag{168}$$

$$=e^{\ell+T_+}\sum_{n=N}^{\infty}e^{(\ell+T_+)n}\exp\left(-\frac{1}{2}e^{(\ell-T_-)n}\Delta\right)= \tag{169}$$

$$=e^{\ell+T_+}\cdot\left(\sum_{n=1}^{\infty}e^{(\ell+T_+)(n+1)}\exp\left(-\frac{1}{2}e^{(\ell-T_-)n}\Delta\right)-\sum_{n=1}^{N-1}e^{(\ell+T_+)(n+1)}\exp\left(-\frac{1}{2}e^{(\ell-T_-)n}\Delta\right)\right). \tag{170}$$

Now a simple computation shows that the function

$$(0,\infty)\ni x\mapsto e^{(\ell+T_+)x}\exp\left(-\frac{1}{2}e^{-(\ell-T_-)x}\Delta\right)\in\mathbb{R} \tag{171}$$

has a single maximum at

$$x^*=\frac{1}{\ell-T_-}\log\left(\frac{2(\ell+T_+)}{\Delta(\ell-T_-)}\right) \tag{172}$$

and that it increases on $(0,x^*)$ and decreases on (x^*,∞). We can thus estimate the sum from 1 to $\lfloor x^*\rfloor$ from above by the integral from 1 to $\lceil x^*\rceil$ and the sum from $\lceil x^*\rceil$ to infinity by the integral from $\lfloor x^*\rfloor$ to infinity. This combined with the fact that the integral over $(\lfloor x^*\rfloor,\lceil x^*\rceil)$ is bounded from above by the value of the function at x^* gives us, after calculating the integral from 0 to infitnity, the following inequality:

$$\sum_{n=1}^{\infty}e^{(\ell+T_+)(n+1)}\exp\left(-\frac{1}{2}e^{(\ell-T_-)n}\Delta\right)\leqslant\Delta^{-L}\cdot\left[2^L\Gamma(L)+(2L)^Le^{-1/4L}\right], \tag{173}$$

where by L we denoted

$$L=\frac{\ell+T_+}{\ell-T_-}. \tag{174}$$

Using this we conclude that

$$\mathsf{brj}^{1[N]}_\omega(\delta) \leqslant e^\delta e^{\ell+T_+}\left(\Delta^{-L}\cdot\left[2^L\Gamma(L)+(2L)^Le^{-1/4L}\right]-\sum_{n=1}^{N-1}e^{(\ell+T_+)(n+1)}\exp\left(-\frac{1}{2}e^{(\ell-T_-)n}\Delta\right)\right). \tag{175}$$

2. Bounds on $\mathsf{brj}^{2[N]}_\omega(\delta)$

$$e^{-\delta}\mathsf{brj}^{2[N]}_\omega(\delta) = \sum_{n=N}^{\infty}\frac{e^{-(q_{n-1}+q_n)\Delta}}{\eta_n}H_{a_{n+1}-1} \leqslant \frac{1}{\log 2}\sum_{n=N}^{\infty}q_{n+1}e^{-(q_{n-1}+q_n)\Delta}\log\frac{q_{n+1}}{q_n} \leqslant \tag{176}$$

$$\leqslant \frac{1}{\log 2}\sum_{n=N}^{\infty}e^{(\ell+T_+)(n+1)}\exp\left(-\frac{1}{2}\Delta\left(1+e^{-(\ell-T_-)}\right)e^{(\ell-T_-)n}\right)\cdot\left[\log 2+\ell+T_++(T_+-T_-)n\right] = \tag{177}$$

$$= \frac{e^{\ell+T_+}}{\log 2}\sum_{n=1}^{\infty}e^{(\ell+T_+)n}\exp\left(-\frac{1}{2}\Delta\left(1+e^{-(\ell-T_-)}\right)e^{(\ell-T_-)n}\right)\cdot\left[\log 2+\ell+T_++(T_+-T_-)n\right] - \tag{178}$$

$$- \frac{e^{\ell+T_+}}{\log 2}\sum_{n=1}^{N-1}e^{(\ell+T_+)n}\exp\left(-\frac{1}{2}\Delta\left(1+e^{-(\ell-T_-)}\right)e^{(\ell-T_-)n}\right)\cdot\left[\log 2+\ell+T_++(T_+-T_-)n\right]. \tag{179}$$

We now use an analogous argument to the one in the previous case to arrive at

$$\frac{1}{\log 2}\sum_{n=1}^{\infty}e^{(\ell+T_+)n}\exp\left(-\frac{1}{2}\Delta\left(1+e^{-(\ell-T_-)}\right)e^{(\ell-T_-)n}\right)\cdot\left[\log 2+\ell+T_++(T_+-T_-)n\right] \leqslant \tag{180}$$

$$\leqslant \left(\frac{\log 2+\ell+T_+}{\log 2}\left[\Gamma(L)\cdot 2^L\left(1+e^{-(\ell-T_-)}\right)^L+(2L)^Le^{-\left(1+e^{-(\ell-T_-)}\right)/4L}\right]+\right. \tag{181}$$

$$\left.+\frac{T_+-T_-}{\log 2}\left[\frac{2^L}{(\ell-T_-)^2\left(1+e^{-(\ell+T_+)}\right)^L}\left(\Gamma'(L)+\Gamma(L)\log\frac{2}{1+e^{-(\ell+T_+)}}+\Gamma(L)\log\frac{1}{\Delta}\right)\right]\right)\cdot\Delta^{-L} \tag{182}$$

and thus altogether

$$\mathsf{brj}^{2[N]}_\omega(\delta) \leqslant e^\delta e^{\ell+T_+}\left(\frac{\log 2+\ell+T_+}{\log 2}\left[\Gamma(L)\cdot 2^L\left(1+e^{-(\ell-T_-)}\right)^L+(2L)^Le^{-\left(1+e^{-(\ell-T_-)}\right)/4L}\right]+\right. \tag{183}$$

$$\left.+\frac{T_+-T_-}{\log 2}\left[\frac{2^L}{(\ell-T_-)^2\left(1+e^{-(\ell+T_+)}\right)^L}\left(\Gamma'(L)+\Gamma(L)\log\frac{2}{1+e^{-(\ell+T_+)}}+\Gamma(L)\log\frac{1}{\Delta}\right)\right]\right)\cdot\Delta^{-L}- \tag{184}$$

$$- \frac{e^\delta e^{\ell+T_+}}{\log 2}\sum_{n=1}^{N-1}e^{(\ell+T_+)n}\exp\left(-\frac{1}{2}\Delta\left(1+e^{-(\ell-T_-)}\right)e^{(\ell-T_-)n}\right)\cdot\left[\log 2+\ell+T_++(T_+-T_-)n\right]. \tag{185}$$

3. Bounds on $\mathsf{brj}^{3[N]}_\omega(\delta)$*.*

$$e^{-\delta}\mathsf{brj}^{3[N]}_\omega(\delta) = \sum_{n=N}^{\infty}\frac{e^{-2q_n\Delta}}{\eta_n}(H_{a_{n+1}}-1) \leqslant \sum_{n=N}^{\infty}q_{n+1}e^{-2q_n\Delta}\log\frac{q_{n+1}}{q_n} \leqslant \tag{186}$$

$$\leqslant \sum_{n=N}^{\infty} e^{(\ell+T_+)(n+1)} \exp\left(-\Delta e^{(\ell-T_-)n}\right) [\log 2 + (T_+ - T_-)n] = \tag{187}$$

$$= e^{\ell+T_+} \sum_{n=1}^{\infty} e^{(\ell+T_+)(n+1)} \exp\left(-\Delta e^{(\ell-T_-)n}\right) [\log 2 + (T_+ - T_-)n] - \tag{188}$$

$$- e^{\ell+T_+} \sum_{n=1}^{N-1} e^{(\ell+T_+)(n+1)} \exp\left(-\Delta e^{(\ell-T_-)n}\right) [\log 2 + (T_+ - T_-)n] \tag{189}$$

We now estimate the sum from 1 to infinity exactly as above and we finally obtain

$$\mathsf{brj}_{\omega}^{3[N]}(\delta) \leqslant e^{\delta} e^{\ell+T_+} \left[\Gamma(L)\log 2 + L^L e^{-1/2L}\log 2 + \frac{T_+ - T_-}{(\ell - T_-)^2}\Gamma'(L) + \frac{T_+ - T_-}{(\ell - T_-)^2}\Gamma(L)\log\frac{1}{\Delta}\right]\Delta^{-L} - \tag{190}$$

$$- \sum_{n=1}^{N-1} e^{(\ell+T_+)(n+1)} \exp\left(-\Delta e^{(\ell-T_-)n}\right) [\log 2 + (T_+ - T_-)n]. \tag{191}$$

We remark that all of the above series are of the order $O(\delta^{-L}\log\delta^{-1})$, which is crucial for the Kolmogorov - Newton iterative scheme.

5 Metric properties of the Khintchine - Lévy condition

In this section we prove that for certain T and N the measure (both Lebesgue and Gauss) of the set $KL^{\circ}(T,N)$ is positive and we also provide lower bounds for it ($\circ \in \{+,-,\ \}$ from here on).

Note that obtaining lower bounds on the measure of $KL^{\circ}(T,N)$ is the same as obtaining upper bounds on the measure of $KL^{\circ}(T,N)^c$ and, according to (27), we have

$$KL^{\circ}(T,N)^c = \bigcup_{n=N}^{\infty} KL_n^{\circ}(T)^c, \tag{192}$$

therefore it is enough to estimate the measure of $KL_n^{\circ}(T)^c$ from above by a quantity which gives a convergent series with sum less than 1. We will indeed show that this upper bound is exponential in $-n$. We begin our considerations with a result related to the Khintchine-Lévy theorem (theorem 2.1).

Lemma 5.1. *The following inequalities hold for all $n \geqslant 1$:*

$$-\log 2 < \mathbb{E}_{\gamma}\log q_n - n\ell < 0. \tag{193}$$

Proof. First observe that

$$\frac{1}{\log 2}\int_0^1 \frac{\log x}{x+1}dx = -\ell, \tag{194}$$

thus, using the invariance of γ under G, we can conclude that for $n = 1, 2, \ldots$

$$\mathbb{E}_{\gamma}\log x_n = \ell. \tag{195}$$

Using this (and lemma 2.1) we see that

$$\mathbb{E}_{\gamma}\log(\eta_{n-1}^{-1}) = \mathbb{E}_{\gamma}\log(x_1 \ldots x_n) = n\ell. \tag{196}$$

This way for our purposes it is enough to estimate $\mathbb{E}_\gamma \log(q_n \eta_{n-1})$. We know, however, in view of inequalities (102) that

$$\frac{1}{2} < q_n \eta_{n-1} < 1 \tag{197}$$

and the result follows. $\square$

Denote

$$X_n = \log \frac{q_n}{q_{n-1}} \tag{198}$$

and

$$S_n = X_1 + \ldots + X_n = \log q_n \tag{199}$$

for $n = 1, 2, \ldots$. Observe that

$$KL_n(T)^c = \left\{\omega \in X : \frac{\log q_n}{n} \geqslant \ell + T\right\} = \left\{\omega \in X : \frac{(X_1 - \ell) + (X_2 - \ell) + \ldots + (X_n - \ell)}{n} \geqslant T\right\} = \tag{200}$$

$$= \left\{\omega \in X : \frac{S_n}{n} - \ell \geqslant T\right\}. \tag{201}$$

This way we can regard the measure of $KL_n(T)^c$ as the tail of the probability distribution of the "almost" centering of random variables S_n/n (the "almost" part is because of lemma 5.1). Now the Khintchine-Lévy theorem (theorem 2.1) tells us that a strong law of large numbers holds for the sequence (X_n). We would now like to obtain a stronger result - a quantitative central limit theorem. The problem with our setting is that the random variables X_j are not independent as required in the classical CLT-like results. They are, however, close to being independent - they satisfy the following condition of *ψ-mixing*:

Definition 5.1. *Let $(Y_n)_{n=1}^{\infty}$ be a sequence of random variables on a probability space $(\mathcal{Y}, \Sigma, \mu)$. For indices $a \leqslant b \in \mathbb{N}_* \cup \{\infty\}$ denote by σ_a^b the σ-algebra generated by random variables Y_ν with $a \leqslant \nu \leqslant b$. We say that the sequence (Y_n) is* ψ-mixing *(w.r.t. μ) if*

$$\sup \left| \frac{\mu(A \cap B)}{\mu(A)\mu(B)} - 1 \right| \overset{n\to\infty}{\longrightarrow} 0, \tag{202}$$

where the supremum is taken over $A \in \sigma_1^k, B \in \sigma_{k+n}^{\infty}$ and $k \in \mathbb{N}_*$.

The ψ-mixing property tells us the following: the more two events A and B are separated in time the more independent they become.

Lemma 5.2. *The sequence (X_n) is ψ-mixing w.r.t. the Gauss measure γ.*

Proof. Since ψ-mixing depends only on the σ-algebras generated by the sequence and

$$\frac{q_j}{q_{j-1}} = [a_j, \ldots, a_1]^{-1} \tag{203}$$

the claim follows from [16, Theorem 1.3.14] and [16, Proposition 1.3.13]. $\square$

We can now apply the results of large deviations theory, which makes it possible to control the tails of probability distributions (and thus gives the desired quantitative CLT-like result) knowing how the moments of X_n behave. We give the flavor of our strategy by the following two theorems taken from [28].

Theorem 5.1 ([28, Theorem 4.21])**.** *Let $(Y_n)_{n=1}^{\infty}$ be a sequence of random variables defined on a probability space $(\mathcal{Y}, \Sigma, \mu)$ and denote $W_n = Y_1 + \ldots + Y_n$. Assume that $Y_n = f_n(\xi_n)$, where f_n is a real valued measurable function and ξ_n is a Markov chain (w.r.t. μ) and that (Y_n) is a ψ-mixing sequence (w.r.t. to μ). Assume also that it satisfies the following moment estimate:*

$$\mathbb{E}_\mu |Y_n|^k \leqslant (k!)^{1+\gamma_1} H_1^k \tag{204}$$

for some constants $\gamma_1 \geqslant 0$ and $H_1 > 0$ and all integers $k \geqslant 2$ and $n \geqslant 1$. Then for each $n \geqslant 1$ the following inequality holds:

$$|\Gamma_k(W_n)| \leqslant (k!)^{1+\gamma_1} 16^{k-1} H_1^k n, \tag{205}$$

where the cumulants Γ_k are taken with respect to μ.

Theorem 5.2 ([28, Lemma 2.4], [4])**.** *Let W be a centered random variable defined on a probability space $(\mathcal{Y}, \Sigma, \mu)$ (i.e. $\mathbb{E}_\mu W = 0$). Assume there exist constants $\gamma_2 \geqslant 0, H > 0$ and $\bar{\Delta} > 0$ such that for all integers $k \geqslant 2$ we have*

$$|\Gamma_k(W)| \leqslant \left(\frac{k!}{2}\right)^{1+\gamma_2} \frac{H}{\bar{\Delta}^{k-2}}. \tag{206}$$

Then for all $x \geqslant 0$ the following inequality is valid:

$$\mu(\pm W \geqslant x) \leqslant \exp\left(- \frac{x^2}{2\left(H + \left(x/\bar{\Delta}^{1/(1+2\gamma_2)}\right)\right)^{(1+2\gamma_2)/(1+\gamma_2)}} \right). \tag{207}$$

Here Γ_k denotes the cumulant taken w.r.t. μ, while the notation $\pm W$ indicates that the inequality holds both for W and $-W$.

In order to be able to apply theorems 5.1 and 5.2 we need to verify if the random variables X_j are associated to a Markov chain.

Lemma 5.3. *The sequence $(q_{n-1}/q_n)_{n=1}^{\infty}$ is a $\mathbb{Q}$-valued Markov chain w.r.t. γ.*

Proof. We will show a much stronger claim - that $s_n = q_{n-1}/q_n$ satisfies the Markov property for any measure, for which the definition of a Markov chain makes sense. In other words we claim that if our sequence at time n is in a state $\bar{s}_n$, then we can retrieve all its past states. Since

$$s_n = [a_n, \ldots, a_1] \tag{208}$$

the state $\bar{s}_n$ must necessarily be a rational number, whose continued fraction expansion has length equal to n, otherwise the set $\{s_n = \bar{s}_n\}$ is empty. In this case it is enough to compute the continued fraction expansion of $\bar{s}_n$ (we denote it by $[\bar{a}_n, \ldots, \bar{a}_1]$) and the past states are uniquely given by

$$\bar{s}_j = [\bar{a}_j, \ldots, \bar{a}_1] \tag{209}$$

since $\bar{s}_n$ has exactly one continued fraction expansion with length equal to n.

In other words the conditional probability of $\{s_n = \bar{s}_n\}$ under the assumption that $\{s_{n-1} = \bar{s}_{n-1}, \ldots, s_1 = \bar{s}_1\}$ is either ill-defined or equal to zero (this happens when one of the equalities $\bar{s}_{n-j} = G^j(\bar{s}_n), j = 0, \ldots, n-1$ is not valid, in particular the second scenario happens when $\bar{s}_{n-j} = G^j(\bar{s})$ with $\bar{s} \neq \bar{s}_n$) or it is equal to the probability of $\{s_n = \bar{s}_n\}$ under $\{s_{n-1} = G(\bar{s}_n), s_{n-2} = G^2(\bar{s}_n), \ldots, s_1 = G^{n-1}(\bar{s}_n)\}$, in which case it is equal to the probability of $\{s_n = \bar{s}_n\}$ under $\{s_{n-1} = G(\bar{s}_n)\}$. □

It is clear that upon substituting $S_n - \mathbb{E}_\gamma S_n$ for W and nT for x in theorem 5.2 we almost obtain desired estimates on $KL_n^+(T)^c$ and we only need to make corrections taking into account lemma 5.1, since $\mathbb{E}_\gamma \log q_n$ is only approximately equal to $n\ell$. We thus only have to verify the assumptions of theorem 5.1 with $Y_n = X_n$ to obtain estimates on k-th cumulants of S_n with $k \geqslant 2$ (which are also estimates on cumulants of $S_n - \mathbb{E}_\gamma S_n$, since centering only affects the first cumulant) to later use them in theorem 5.2.

Theorem 5.3. *The Gauss measure of $KL(T_-, T_+, N)^c$ satisfies*

$$\gamma(KL(T_-, T_+, N)^c) \leqslant \beta(T_-, N) + \beta(T_+, N), \tag{210}$$

where

$$\beta(T, N) = \frac{\exp\left(-\frac{T^2}{32\bar{r}(2\bar{r}+T)}N\right)}{1 - \exp\left(-\frac{T^2}{32\bar{r}(2\bar{r}+T)}\right)} \tag{211}$$

and

$$\bar{r} = \sup_{k \geqslant 2} \zeta(k+1)^{1/k} = \sqrt{\zeta(3)} < 1.097. \tag{212}$$

Proof. We first give an upper bound of the measure of $KL^+(T, N)^c$. We begin with estimating the k-th moment of X_n for $k \geqslant 2$. We have

$$\mathbb{E}_\gamma |X_n|^k = \mathbb{E}_\gamma \left|\log \frac{q_n}{q_{n-1}}\right|^k \leqslant \mathbb{E}_\gamma |\log(1 + a_n)|^k \overset{(\star)}{=} \mathbb{E}_\gamma |\log(1 + a_1)|^k = \int_0^1 \frac{|\log(1 + \lfloor x^{-1} \rfloor)|^k}{1+x} dx \leqslant \tag{213}$$

$$\leqslant \int_0^1 \frac{|\log(1 + x^{-1})|^k}{1+x} dx = \int_1^\infty \frac{\log(1+y)^k}{y^2+y} dy = \int_{\log 2}^\infty \frac{z^k e^{-z}}{1 - e^{-z}} dz = \tag{214}$$

$$= \sum_{j=1}^\infty \int_{\log 2}^\infty z^k e^{-jz} dz = \sum_{j=1}^\infty \frac{1}{j^{k+1}} \int_{\log 2/j}^\infty u^k e^{-u} du \leqslant \zeta(k+1) \cdot k! \leqslant \bar{r}^k \cdot k!. \tag{215}$$

Equality $(\star)$ is a consequence of G-invariance of γ, while in the following equalities we simply substituted x for $1/y$, $1+y$ for e^z and jz for u, respectively. The exponential estimate of the Riemann zeta function is much of an overkill of course, however we introduced it so that the inequality fits the framework of theorem 5.1. Together with the bounds derived above lemmas 5.3 and 5.2 tell us, that the assumptions of theorem 5.1 are indeed satisfied, thus we have

$$|\Gamma_k(S_n)| \leqslant k! \cdot (16\bar{r})^{k-2} \cdot 16\bar{r}^2 n \tag{216}$$

for $k \geqslant 2$. However, since shifting a random variable by a constant affects only its first cumulant we also have

$$|\Gamma_k(S_n - \mathbb{E}_\gamma S_n)| \leqslant k! \cdot (16\bar{r})^{k-2} \cdot 16\bar{r}^2 n. \tag{217}$$

This allows us to apply theorem 5.2 with $\gamma_2 = 0$, $\bar{\Delta} = (16\bar{r})^{-1}$ and $H = 32\bar{r}^2 n$, which gives

$$\gamma\left(S_n - \mathbb{E}_\gamma S_n \geqslant x\right) \leqslant \exp\left(-\frac{x^2}{2(32\bar{r}^2 n + 16\bar{r}x)}\right). \tag{218}$$

Observe now that

$$\{\omega : S_n - n\ell \geqslant x\} \subset \{\omega : S_n - \mathbb{E}_\gamma S_n \geqslant x\} \tag{219}$$

thanks to lemma 5.1, which after setting $x = nT_+$ implies

$$\gamma(S_n - n\ell \geqslant nT_+) \leqslant \exp\left(-\frac{T_+^2}{32\bar{r}(2\bar{r} + T_+)} n\right) \tag{220}$$

which is equivalent to

$$\gamma(KL_n^+(T_+)^c) \leqslant \exp\left(-\frac{T_+^2}{32\bar{r}(2\bar{r}+T_+)}n\right). \tag{221}$$

For a given natural number N we then have

$$\gamma(KL^+(T_+,N)^c) \leqslant \sum_{n=N}^{\infty} \exp\left(-\frac{T_+^2}{32\bar{r}(2\bar{r}+T_+)}n\right) = \frac{\exp\left(-\frac{T_+^2}{32\bar{r}(2\bar{r}+T_+)}N\right)}{1-\exp\left(-\frac{T_+^2}{32\bar{r}(2\bar{r}+T_+)}\right)} = \beta(T_+,N). \tag{222}$$

The reasoning in the case of $KL^-(T_-,N)^c$ is completely analogous. We first use theorem 5.2 (the $-W$ version) to arrive at

$$\gamma\left(S_n - \mathbb{E}_\gamma S_n \leqslant -x\right) \leqslant \exp\left(-\frac{x^2}{2(32\bar{r}^2 n + 16\bar{r}x)}\right). \tag{223}$$

Thanks to lemma 5.1 we conclude that

$$\{\omega : S_n - n\ell + \log 2 \leqslant -x\} \subset \{\omega : S_n - \mathbb{E}_\gamma S_n \leqslant -x\}, \tag{224}$$

which after setting $x = nT_-$ brings us to

$$\gamma(S_n - n\ell + \log 2 \leqslant -nT_-) \leqslant \exp\left(-\frac{T_-^2}{32\bar{r}(2\bar{r}+T_-)}n\right), \tag{225}$$

which is equivalent to

$$\gamma(KL_n^-(T_-)^c) \leqslant \exp\left(-\frac{T_-^2}{32\bar{r}(2\bar{r}+T_-)}n\right) \tag{226}$$

(notice how $\log 2$ generates the $1/2$ factor present in the definition of KL^-). Taking the sum over $n \geqslant N$ allows us to conclude that

$$\gamma(KL^-(T_-,N)^c) \leqslant \beta(T_-,N) \tag{227}$$

and thus

$$\gamma(KL(T_-,T_+,N)^c) \leqslant \gamma(KL^-(T_-,N)^c) + \gamma(KL^+(T_+,N)^c) \leqslant \beta(T_-,N) + \beta(T_+,N). \tag{228}$$

□

Corollary 5.4. *Results of theorem 5.3 hold* mutatis mutandis *for the Lebesgue measure as it is equivalent to the Gauss measure with estimates*

$$\frac{1}{2\log 2}\lambda(B) \leqslant \gamma(B) \leqslant \frac{1}{\log 2}\lambda(B) \tag{229}$$

valid for any Lebesgue measurable set B.

Corollary 5.5. *For $D > 1$ the inequality*

$$\gamma(KL(T,N)) \geqslant 1 - \frac{1}{D} \tag{230}$$

holds whenever

$$N > \frac{32\bar{r}(T+2\bar{r})}{T^2}\left(\log\frac{32\bar{r}(T+2\bar{r})}{T^2} + \log 2D\right), \tag{231}$$

and

$$\gamma(KL^+(T,N)) \geqslant 1 - \frac{1}{D} \tag{232}$$

holds whenever

$$N > \frac{32\bar{r}(T+2\bar{r})}{T^2}\left(\log\frac{32\bar{r}(T+2\bar{r})}{T^2} + \log D\right). \tag{233}$$

In particular we have e.g.

$$\gamma(KL(0.3, 9607)) \geqslant 0.9, \quad \gamma(KL(0.3, 11847)) \geqslant 0.99, \quad \gamma(KL(0.3, 14087)) \geqslant 0.999 \tag{234}$$

and

$$\gamma(KL^+(2, 218)) \geqslant 0.9, \quad \gamma(KL^+(2, 303)) \geqslant 0.99, \quad \gamma(KL^+(2, 387)) \geqslant 0.999. \tag{235}$$

References

[1] V. I. Arnol'd. Small denominators. I. Mapping the circle onto itself. *Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya*, 25:21–86. English translation in *Amer. Math. Soc. Transl. (2),* 46:213–284, 1965.

[2] V. I. Arnol'd. Proof of a theorem of A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian. *Russian Mathematical Surveys*, 9:9–36, 1963.

[3] Arnol'd, V. I. *Geometrical Methods In The Theory Of Ordinary Differential Equations.* Springer New York, 1988.

[4] R. Bentkus and R. Rudzkis. On exponential estimates of the distributions of random variables. *Lithuanian Mathematical Journal*, 20:15–30, 1980.

[5] H. W. Broer. KAM theory: The legacy of Kolmogorov's 1954 paper. *Bulletin of the American Mathematical Society*, 41, 2004.

[6] A. Celletti and L. Chierchia. Construction of analytic KAM surfaces and effective stability bounds. *Communications in Mathematical Physics*, 161:119–161, 1988.

[7] A. Celletti and L. Chierchia. A constructive theory of Lagrangian tori and computer-assisted applications. *Dynamics Reported*, 4:60–129, 1995.

[8] P. Cvitanović. Circle maps: irrationally winding. In C. Itzykson, J.-M. Luck, P. Moussa, and Waldschmidt M., editors, *From Number Theory to Physics.* Springer Berlin Heidelberg, 1992.

[9] R. de la Llave. A simple proof of a particular case of C. Siegel's center theorem. *Journal of Mathematical Physics*, 24(8):2118–2121, 1983.

[10] R. de la Llave. A tutorial on KAM theory. In R. de la Llave, A. Katok, Ya. Pesin, and H. Weiss, editors, *Smooth Ergodic Theory and its Applications*, Proceedings of Symposia in Pure Mathematics, 2000.

[11] R. de la Llave, A. González, À. Jorba, and J. Villanueva. KAM theory without action-angle variables. *Nonlinearity*, 18:855–895, 2005.

[12] R. de la Llave and D. Rana. Accurate strategies for small divisor problems. (preprint, pages 1-87).

[13] R. de la Llave and D. Rana. Accurate strategies for small divisor problems. *Bulletin (New Series) of the American Mathematical Society*, 22(1):85–90, 1990.

[14] M. Herman. Sur les courbes invariantes par le difféomorphismes de l'anneau. *Asterisque*, 144, 1986.

[15] I. A. Ibragimov. A theorem from the metric theory of continued fractions. *Vestnik Leningrad. Univ.*, 16(1):13–24, 1961.

[16] M. Iosifescu and C. Kraaikamp. *Metrical Theory of Continued Fractions*. Springer Netherlands, 2002.

[17] A. N. Kolmogorov. On conservation of conditionally periodic motions for a small change in Hamilton's function. *Doklady Akademii Nauk SSSR (N.S.)*, 98:527–530. English translation in *Stochastic Behavior in Classical and Quantum Hamiltonian Systems (Volta Memorial Conf., Como, 1977)*, Lecture Notes in Phys., 93, pages 51–56. Springer, Berlin, 1979.

[18] P. Lévy. Sur le développement en fraction continue d'un nombre choisi au hasard. *Compositio Math.*, 3:286–303, 1936.

[19] S. Marmi. An introduction to small divisors. arXiv:math/0009232, 2000.

[20] S. Marmi. From small divisors to Brjuno functions (New Trends and Applications of Complex Asymptotic Analysis: around dynamical systems, summability, continued fractions). *RIMS Kokyuroku*, 1493:1–47, 2006.

[21] G. Misevičius. Estimate of the remainder term in the limit theorem for denominators of continued fractions. *Lithuanian Mathematical Journal*, 21(3):245–253, 1981.

[22] J. K. Moser. On Invariant Curves of Area Preserving Mappings of an Annulus. *Nachrichten der Akademie der Wissenschaften in Göttingen: II, Mathematisch-Physikalische Klasse*, K1:1–20, 1962.

[23] W. Philipp and O. P. Stackelberg. Zwei Grenzwertsätze für Kettenbrüche. *Mathematische Annalen*, 2:152–156, 1969.

[24] J. Pöschel. Integrability of Hamiltonian systems on Cantor sets. *Communications on Pure and Applied Mathematics*, 35:653–696, 1982.

[25] H. Rüssmann. On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus. In *Dynamical Systems, Theory and Applications*, pages 598–624. Springer Berlin Heidelberg, 1975.

[26] H. Rüssmann. On optimal estimates for the solutions of linear difference equations on the circle. *Celestial Mechanics*, 14(1):33–37, 1976.

[27] H. Rüssmann. Stability of elliptic fixed points of analytic area-preserving mappings under the Bruno condition. *Ergodic Theory and Dynamical Systems*, 22(05):1551–1573, 2002.

[28] L. Saulis and V. A. Statulevičius. *Limit theorems for large deviations*. Kluwer Academic Publishers, 1991.

[29] C. L. Siegel. Iteration of analytic functions. *Annals of Mathematics*, 43:607–612, 1942.

[30] J.-C. Yoccoz. Introduction to small divisors problems. In C. Itzykson, J.-M. Luck, P. Moussa, and M. Waldschmidt, editors, *From Number Theory to Physics*. Springer Berlin Heidelberg, 1992.

[31] J.-C. Yoccoz. Analytic linearization of circle diffeomorphisms. In *Dynamical Systems and Small Divisors*, pages 125 – 173. Springer Berlin Heidelberg, 2002.