STRONG UNIFORM DISTRIBUTION — THE CASE OF INFINITELY MANY PRIMES

GADY KOZMA

ABSTRACT. We show that for any infinite set of primes $\mathcal{P}$, the series of all multiples of powers of primes from $\mathcal{P}$ is not L^∞-good.

1. Introduction

One of the starting points for this topic is the following result of Marstrand, [M70], answering a question of Khinchin [K23]:

Marstrand's theorem. *There exists a function $f \in L^\infty([0,1])$ such that the limit*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{r=1}^{n} f(\langle rx \rangle) \tag{1}$$

does not exist almost everywhere, where $\langle y \rangle$ is the fractional part of y i.e. $y - \lfloor y \rfloor$.

While the formulation is very elegant, it really hides the fact that in a certain sense this is an infinite dimensional ergodic theorem. To see this, replace "r" in (1) with "p^r" and you will get an ergodic question (for the transformation $x \mapsto \langle px \rangle$), and in effect, in this case the

2000 *Mathematics Subject Classification.* 11K06, 37A45.
Key words and phrases. Strong uniform distribution, Infinitely many primes, L^∞-good series.
This work is part of the research program of the European Network "Analysis and Operators", contract HPRN-CT-00116-2000 supported by the European Commission.

sum converges for every $f \in L^\infty$ to $\int f(t)\,dt$ [R36]. This one dimensional case has been researched extensively. See e.g. [B88, B89, H92]. Marstrand concludes from Raikov's result [R36] that the same is true in the finite dimensional case, i.e. for the sequence $a(\mathcal{P})$ (defined below) where $\mathcal{P}$ is finite set of primes.

Definition. For $\mathcal{P}$ a set of primes we define the sequence

$$a(\mathcal{P}) := \{a \geq 1 \,:\, p|a \Rightarrow p \in \mathcal{P}\} \quad .$$

Here, and everywhere, since all the sequences will be increasing we shall not distinguish between a sequence a_r and the set $\{a_r\}_{r=1}^\infty$, and will freely use set notations such as $n \in a(\mathcal{P})$and $a(\mathcal{P}_1) \setminus a(\mathcal{P}_2)$.

This last result of Marstrand has been generalized by R. Nair [N90] from L^∞ to L^1. Thus it became easier to use the following language:

Definition. For an increasing sequence of integers a_r and a class of functions $\mathcal{A}$ we say that a_r is an $\mathcal{A}$-good sequence (called an $\mathcal{A}^*$ sequence in [N90, N01]) if for every $f \in \mathcal{A}$ we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{r=1}^{n} f(\langle a_r x\rangle) = \int_0^1 f(t)\,dt$$

almost everywhere with respect to the Lebesgue measure.

With this formulation, Marstrand's result is that $a_r = r$ is not L^∞-good while $a(\mathcal{P})$ is L^∞-good for every finite $\mathcal{P}$. Nair's result is that $a(\mathcal{P})$ is L^1-good, and in the same paper [N90] he asks: is the condition that $\mathcal{P}$ is finite necessary? The main result of this paper is the answer[1] to that question:

Theorem 1. *For every infinite set of primes $\mathcal{P}$, the sequence $a(\mathcal{P})$ is not L^∞-good.*

Note that a false proof of the opposite of theorem 1 was published in [L98]. You do not have to take my word for it, though. For example, [N01] mentions in passing that the proof in [L98] is false.

This paper is organized as follows: the proof of theorem 1 can be found in chapter 2, relying heavily on techniques from [M70]. Chapter 3 sketches a generalization to infinitely generated multiplicative semigroups of integers.

2. Infinitely many primes

We shall require the following lemmas from [M70]:

Lemma 1. *Suppose that for every q, $v \geq 1$ one can find sets of positive integers G, H with the properties*

(1) $|G| > v|H|$

(2) *for every $g \in G$ there exists $\eta \geq 1$ such that*

$$\frac{g}{a_r} \in H \quad \forall \eta \leq r \leq \eta q$$

then the sequence a_r is not L^∞-good.

This is corollary 3.3 from [M70].

[1] Indeed, a generalized answer, since not being L^∞-good is obviously stronger than not being L^1-good.

Lemma 2. *For every positive integer s and any primes $p_1, \ldots, p_s$,*

$$\#(a(p_1, \ldots, p_s) \cap [1, x]) \sim \left(\frac{1}{s!} \prod_{j=1}^{s} \frac{1}{\log p_j} \right) (\log x)^s$$

where $f(x) \sim g(x)$ is short for

$$\lim_{x \to \infty} \frac{f(x)}{g(x)} = 1 \quad .$$

This is lemma 4.1 from [M70]. Marstrand proves it only for $\{p_1, \ldots, p_s\}$ being the first s primes, but the proof carries over to the general case without any change. Roughly, Marstrand's proof consists of taking log and then estimating the intersection of a lattice with a simplex by the volume of the simplex.

Lemma 3. *Let s be a positive integer $s \geq 2$; let $p_1, \ldots, p_s$ be the first s primes; and let $q > 1$ be any real number. Than*

$$\#(a(p_1, \ldots, p_s) \cap [x, qx]) \sim \left(\frac{1}{(s-1)!} \prod_{j=1}^{s} \frac{1}{\log p_j} \right) \log q (\log x)^{s-1}$$

as $x \to \infty$.

This is lemma 4.2 from [M70]. It is not difficult to see that here too the set of primes need not be limited, but we shall have no use for this fact (despite their similarity, these two lemmas are used in rather different ways).

Definition. For a sequence $a = \{a_r\}_{r=1}^{\infty}$ and integers q and h we shall denote by $A_{h,q}(a)$ the number

$$A_{h,q}(a) := \max_{a_r \leq h} \frac{a_{qr}}{a_r} \quad . \tag{2}$$

If $a = a(\mathcal{P})$ we shall write for short $A_{h,q}(\mathcal{P}) := A_{h,q}(a(\mathcal{P}))$.

Lemma 4. *If* $\mathcal{P}_1 \subset \mathcal{P}_2$ *then*

$$A_{h,q}(\mathcal{P}_1) \geq A_{h,q}(\mathcal{P}_2) \quad .$$

Proof. For $i = 1, 2$ denote $\{a^i_r\}_{r=1}^{\infty} := a(\mathcal{P}_i)$ and $A^i_{h,q} = A_{h,q}(\mathcal{P}_i)$. We note that for any particular h the question involves only a finite set of primes, therefore without loss of generality we may assume $\mathcal{P}_2 = \mathcal{P}_1 \cup \{p\}$. Let r satisfy $a^2_r \leq h$. The lemma will be proved once we show that

$$\frac{a^2_{qr}}{a^2_r} \leq A^1_{h,q} \quad .$$

Define s using

$$a^1_s \leq a^2_r < a^1_{s+1} \tag{3}$$

Examining $(a^2 \setminus a^1) \cap [1, a^2_r]$ we find that it is a union of n disjoint sets $S_1, \ldots, S_n$ each of which is a copy of (a part of) a^1 multiplied by p^k i.e.

$$S_k = \{p^k a^1_1, ..., p^k a^1_{r(k)}\} \quad .$$

Here n is given by

$$p^n \leq a^2_r < p^{n+1}$$

and $r(k)$ by

$$p^k a^1_{r(k)} \leq a^2_r < p^k a^1_{r(k)+1} \tag{4}$$

and these give $r = s + \sum r(k)$.

We have two cases to examine. First, assume that for some k we have $a^2_{qr} \le p^k a^1_{qr(k)}$. In this case (4) and (2) give

$$\frac{a^2_{qr}}{a^2_r} \le \frac{p^k a^1_{qr(k)}}{p^k a^1_{r(k)}} \le A^1_{h,q} \quad .$$

Therefore we are left with the second case, which is $a^2_{qr} > p^k a^1_{qr(k)}$ for all k. In this case we use the obvious fact that, if $M := \#\left((a^2 \setminus a^1) \cap \left[1, a^2_j\right[\right)$ then

$$a^2_j \le a^1_{j-M} \quad .$$

Our assumption $a^2_{qr} > p^k a^1_{qr(k)}$ for all k gives $M \ge q \sum r(k)$ and then

$$a^2_{qr} \le a^1_{qr-M} \le a^1_{qr-q\sum r(k)} = a^1_{qs}$$

So with (3) we again get

$$\frac{a^2_{qr}}{a^2_r} \le \frac{a^1_{qs}}{a^1_s} \le A^1_{h,q} \quad . \qquad \square$$

Lemma 5. *If $\mathcal{P}$ is infinite and q is any integer then*

$$\lim_{h\to\infty} \frac{\log h}{\log A_{h,q}(\mathcal{P})} = \infty \quad .$$

Proof. We first calculate this quantity for a finite set of primes $\mathcal{Q}$. Assume $|\mathcal{Q}| = s$. Then by lemma 2 we have

$$\log h \sim K r^{1/s} \quad \forall a_r \le h < a_{r+1}$$

for some constant K, as $h \to \infty$. This gives

$$\log \frac{a_{qr}}{a_r} \sim K r^{1/s} \left(q^{1/s} - 1\right) \quad . \tag{5}$$

From this it is easy to get an estimate for $\log A_{h,q}$ since the fact that the right hand side of (5) is a rising function of r allows to estimate the maximum in (2) by the largest r such that $a_r \leq h$ and we get

$$\log A_{h,q} \sim K r^{1/s}(q^{1/s} - 1)$$

so

$$\lim_{h\to\infty} \frac{\log h}{\log A_{h,q}(\mathcal{Q})} = \frac{1}{q^{1/s} - 1} \quad .$$

This, together with lemma 4 finishes the proof. □

Proof of theorem 1. Since my vehicle is lemma 1, let q and v be two integers. Take p to be some arbitrary prime in $\mathcal{P}$. Let R be some number sufficiently large satisfying

$$\frac{\log \frac{a_R}{a_q}}{\log p A_{a_R,q}} > v + 1 \tag{6}$$

which is clearly possible, using lemma 5. Further, let p_s be the largest prime dividing one of $a_1, ..., a_{qR}$, define the sequence $b = a(p_1, ..., p_s)$ (i.e. all the primes up to p_s) and let $m = \prod_{j=1}^{qR} a_j$ and $A = pA_{a_R,q}$. For an as yet unspecified x define the sets

$$\begin{aligned} H &= [x, Ax] \cap b \\ G &= [a_q x, A a_R x] \cap mb \end{aligned}$$

where as usual $mb = \{mb_r : b_r \in b\}$.

We need to verify the two conditions of lemma 1, and we start with the sizes of H and G. Lemma 3 gives for some constant K (which depends on s but not on x), and for x sufficiently large,

$$\begin{aligned} |H| &\sim K\log A(\log x)^{s-1} \\ |G| &\sim K\log\frac{Aa_R}{a_q}\left(\log\frac{x}{m}\right)^{s-1} \end{aligned}$$

which clearly implies (remember (6)) that

$$\frac{|G|}{|H|} \geq \frac{\log\frac{a_R}{a_q}}{\log A}(1+o(1)) \geq v+1+o(1)$$

so for x sufficiently large condition 1 will be satisfied.

To check condition 2 of lemma 1, examine for some $\eta \leq R$ the set

$$G_\eta := \left\{ g \in mb \, : \, \frac{g}{a_r} \in H \quad \forall \eta \leq r \leq q\eta \right\} \quad .$$

$g \in mb$ implies $\frac{g}{a_r} \in b$ so from the definition of H we get

$$G_\eta = [a_{q\eta}x, a_\eta Ax] \cap mb \quad .$$

It is for this point that we defined $A_{h,q}$ since we can now write, for $\eta < R$,

$$a_{q(\eta+1)} \leq a_{\eta+1}A_{a_R,q} \leq a_\eta A$$

and thus the intervals $[a_{q\eta}x, a_\eta Ax]$ and $[a_{q(\eta+1)}x, a_{\eta+1}Ax]$ intersect, which means that their union is an interval, and that leads inductively to

$$\bigcup_{\eta=1}^{R} G_\eta = [a_q x, a_R Ax] \cap mb = G$$

and condition 2, and the theorem, are proved. □

Remark. In essence, this is a theorem about densities — it shows that if a sequence is denser than $e^{r^{1/s}}$ for all s then it is not L^∞-good, and the relation between a_r and a_{qr} can be thought of as some kind of regularity condition. Thus it improves on Marstrand's polynomial condition. For example, the same proof shows that if $a_r \sim e^{\log^k r}$ for some k then a_r is not L^∞-good. In the other direction it is not possible to get a result such as "any sequence sufficiently sparse is L^∞-good". Indeed, Marstrand himself shows that any sequence of pairwise coprime integers, no matter how sparse, is not L^∞-good.

3. Infinitely generated semigroups

One might hope to strengthen theorem 1 for any infinitely generated semigroup of integers, but this is not true. For example, the semigroup $\mathcal{S}$ generated by

$$\left\{2^n 3^{\lfloor\sqrt{n}\rfloor},\, 2^{\lfloor\sqrt{n}\rfloor}3^n\right\}_{n=1}^{\infty}$$

($\lfloor x \rfloor$ denoting the integer value of x) is infinitely generated but is almost equivalent to $a(2,3)$, that is

$$\lim_{x\to\infty} \frac{|[1,x]\cap\mathcal{S}|}{|[1,x]\cap a(2,3)|} = 1$$

and since $a(2,3)$ is L^∞-good, so is $\mathcal{S}$. Thus we restrict our attention to semigroups not contained in any $a(\mathcal{Q})$, $\mathcal{Q}$ finite. Namely we want to prove

Theorem 2. *Any multiplicative semigroup of integers $\mathcal{S}$ not contained in any $a(\mathcal{Q})$ for any finite set of primes $\mathcal{Q}$ is not L^∞-good.*

The proof of theorem 2 is similar to that of theorem 1, and therefore I shall only sketch it roughly. We start from lemma 2. The version of lemma 2 for semigroups requires the following procedure. Let $\mathcal{S}$ be a finitely generated multiplicative semigroup of integers. We denote by $p_1, \ldots, p_s$ all the primes participating in $\mathcal{S}$ and write

$$T := \left\{ (\alpha_1, \ldots, \alpha_s) \, : \, \prod p_i^{\alpha_i} \in \mathcal{S} \right\}$$

which is an additive semigroup $\subset \mathbb{Z}^s$. We extend it to a group G and call the dimension of G the basis size related to $\mathcal{S}$.

Lemma 6. *Let $\mathcal{S}$ be a finitely generated multiplicative semigroup of integers and let t be the related basis size. Then there exists a constant K such that*

$$|\mathcal{S} \cap [1, x]| \sim K (\log x)^t \quad .$$

The proof of this lemma is similar to that of lemma 4.1 from [M70], which is, as already remarked, taking logarithms and then estimating the number of lattice point inside a simplex using its volume.

As for the rest of the proof of the theorem, the equivalent of lemma 4 is proved in a similar manner, with the sole change being that (if $\mathcal{S}_2$ is a semigroup spanned by $\mathcal{S}_1$ and m) we must take only such k's satisfying $m^k \notin \mathcal{S}_1$. Lemma 5 is where we use the fact that $\mathcal{S} \not\subset a(\mathcal{Q})$ to get that $\mathcal{S}$ contains sub-semigroups of unbounded related basis size. The proof of the theorem proper is unchanged. $\square$

References

[B88] J. Bourgain, *On maximal ergodic theory for certain subsets of the integers,*Israel Journal of Mathematics 61 (1988), 39–72.

[B89] J. Bourgain, *Pointwise ergodic theorems for arithmetic sets*, Publ. Math. IHES 69 (1989), 5–45.

[H92] Y. Huang, *Random sets for the pointwise ergodic theorem*, Ergodic theory and dynamical systems 12 (1992), 85–94.

[K23] A. Khinchin, *Ein Satz Über Kettenbrüche mit arithmetischen Anwendungen*, Math. Zeitschr. 18 (1923), 289–306.

[L98] Y. Lacroix, *On strong uniform distribution II. The infinite dimensional case*, Acta. Arith. 84 (1998), 279–290.

[M70] J. M. Marstrand, *On Khinchin's conjecture about strong uniform distributions*, Proceeding of the London Mathematical Society 3/21:2 (1970), 540–556.

[N90] R. Nair, *On strong uniform distribution*, Acta Arith. 56 (1990), 183–193.

[N01] R. Nair, *On strong uniform distribution II,* Monatshefte für Mathematik, 132:4 (2001), 341–348.

[R36] D. A. Raikov, *On some arithmetical properties of summable functions*, Mat. Sb. 1/43 (1936), 377–384.

E-mail address: gadykozma@hotmail.com