ON A PAPER OF ERDÖS AND SZEKERES

J. BOURGAIN AND M.-C. CHANG

ABSTRACT. Propositions 1.1 – 1.3 stated below contribute to results and certain problems considered in [E-S], on the behavior of products $\prod_1^n (1 - z^{a_j}), 1 \leq a_1 \leq \cdots \leq a_n$ integers. In the discussion below, $\{a_1, \ldots, a_n\}$ will be either a proportional subset of $\{1, \ldots, n\}$ or a set of large arithmetic diameter.

1. **Introduction**

The aim of this paper is to revisit some of the questions put forward in the paper [E-S] of Erdos and Szekeres.

Following [E-S], define

$$M(a_1, \ldots, a_n) = \max_{|z|=1} \prod_{i=1}^{n} |1 - z^{a_i}| \tag{1.1}$$

where we assume $a_1 \leq a_2 \leq \cdots \leq a_n$ positive integers (in this paper, we restrict ourselves to distinct integers $a_1 < \cdots < a_n$).

Denote

$$f(n) = \min_{a_1 \leq \cdots \leq a_n} M(a_1, \ldots, a_n) \quad \text{and} \quad f_*(n) = \min_{a_1 < \cdots < a_n} M(a_1, \ldots, a_n). \tag{1.2}$$

It was proven in [E-S] that

$$f(n) \geq \sqrt{2n}. \tag{1.3}$$

This lower bound remains presently still unimproved.

In the other direction, [E-S] establish an upper bound

$$f(n) < \exp(n^{1-c}) \text{ for some } c > 0. \tag{1.4}$$

Subsequent improvements were given by Atkinson [A]

$$f(n) = \exp\{O(n^{\frac{1}{2}} \log n)\} \tag{1.5}$$

and Odlyzko [O]

$$f(n) = \exp\{O(n^{\frac{1}{3}}(\log n)^{4/3})\}. \tag{1.6}$$

Also to be mentioned is a construction due to Kolountzakis ([Kol2], [Kol4]) of a sequence $1 < a_1 < \cdots < a_n < 2n + O(\sqrt{n})$ for which

$$f_*(n) \leq M(a_1, \ldots, a_n) < \exp\{O(n^{\frac{1}{2}} \log n)\} \tag{1.7}$$

(Note that Odlyzko's construction does not come with distinct frequencies).

As shown by Atkinson [A], there is a relation between the [E-S] problem and the *cosine-minimum problem*.

Define

$$M_2(n) = \inf\{-\min_\theta \sum_{j=1}^n \cos a_j\theta\} \tag{1.8}$$

with infinum taken over integer sets $a_1 < \cdots < a_n$.

Then

$$\log f_*(n) < O(M_2(n) \log n). \tag{1.9}$$

The problem of determining $M_2(n)$ was put forward by Ankeny and Chowla [C1] motivated by questions on zeta functions.

It is known that $M_2(n) = O(n^{\frac{1}{2}})$ and conjectured by Chowla that in fact $M_2(n) \sim n^{\frac{1}{2}}$ [C2]. The current best lower bound is due to Ruzsa [R]

$$M_2(n) > \exp(c\sqrt{\log n}) \tag{1.10}$$

for some $c > 0$.

As pointed out in [O], polynomials of the form (1.1) are also of interest in connection to Schinzel's problem [S] of bounding the number of irreducible factors of a polynomial on the unit circle in terms of its degree and L^2-norm.

Propositions 1.1 and 1.2 in this paper establish new results for 'dense' sets $S = \{a_1 < \cdots < a_n\}$. The former improves upon (1.7).

Proposition 1.1. *There is a subset* $\{a_1 < \cdots < a_n\} \subset \{1, \ldots, N\}, n \asymp \frac{N}{2}$*, such that*

$$M(a_1, \ldots, a_n) < \exp(c\sqrt{n}\sqrt{\log n} \log\log n). \tag{1.11}$$

On the other hand, the following holds

Proposition 1.2. *There is a constant $\tau > 0$ such that if $\{a_1 < \ldots < a_n\} \subset \{1, \ldots, N\}$ and $n > (1-\tau)N$, then*

$$M(a_1, \ldots, a_n) > \exp \tau n. \tag{1.12}$$

The latter result generalizes the comment made in [E-S] that

$$\lim_{n\to\infty} [M(1, 2, \ldots, n)]^{1/n} \tag{1.13}$$

exists and is between 1 and 2.

In converse direction, one may prove new lower bounds on $M(a_1, \ldots, a_n)$ assuming that the set $\{a_1 < \cdots < a_n\}$ has a sufficiently large arithmetic diameter.

First, we are recalling the notion of a '*dissociated set*' of integers. We say that $D = \{\nu_1, \ldots, \nu_m\} \subset \mathbb{Z}$ is dissociated provided D does not admit non-trivial $0, 1, -1$ relations. Thus

$$\varepsilon_1\nu_1 + \cdots + \varepsilon_m\nu_m = 0 \text{ with } \varepsilon_1 = 0, 1, -1 \tag{1.14}$$

implies

$$\varepsilon_1 = \cdots = \varepsilon_m = 0.$$

A more detailed discussion of this notion and its relation to lacunarity appears in §5 of the paper.

Proposition 1.3. *Assume $\{a_1 < \cdots < a_n\}$ contains a dissociated set of size m. Then*

$$\log M(a_1, \ldots, a_n) \gg \frac{m^{\frac{1}{2}-\varepsilon}}{(\log n)^{1/2}}. \tag{1.15}$$

Hence (1.15) improves upon (1.3) as soon as

$$m \gg (\log n)^{3+\varepsilon}. \tag{1.16}$$

2. **Preliminary estimates**

Let

$$z = e(\theta) = e^{2\pi i\theta}.$$

By taking the real part of $\mathrm{Log}(1-\mathrm{e}^{2\pi\mathrm{i}\theta}) = -\sum_{\mathrm{k}=1}^{\infty}\frac{1}{\mathrm{k}}\mathrm{e}^{2\pi\mathrm{ik}\theta}$, we have

$$\log|1-z| = -\sum_{k=1}^{\infty}\frac{\cos 2\pi k\theta}{k}.$$

Therefore, we have

Fact 1.

$$\prod_{j=1}^{n}|1-z^{a_j}| = e^{-\sum_{j=1}^{n}\sum_{k=1}^{\infty}\ \frac{\cos 2\pi k a_j\theta}{k}}.$$

We first establish some preliminary inequalities for later use.

Since the function e^x is convex, we obtain for any probability measure μ on $\mathbb{T}$ that

$$\prod_{j=1}^{n}|1-e(a_j\theta)| * \mu \geq e^{-(\sum_{j=1}^{n}\sum_{k=1}^{\infty}\ \frac{\cos 2\pi k a_j\ \cdot}{k})*\mu(\theta)}$$

and therefore we have

Fact 2.

$$\Big\|\prod_{j=1}^{n}|1-e(a_j\theta)|\Big\|_{\infty} \geq e^{-\min\limits_{\theta}\{\sum_{j=1}^{n}\sum_{k=1}^{\infty}\frac{\cos 2\pi k a_j\ \cdot}{k}*\mu\}(\theta)}.$$

Lemma 2.1.

$$\log|1-e^{2\pi i\theta}| \leq -\sum_{j=1}^{J}\frac{\rho^j}{j}\cos 2\pi j\theta + O\Big(\frac{1}{\sqrt{J}}\Big) \tag{2.1}$$

where $\rho = 1-\frac{1}{\sqrt{J}}$ *and (2.1) is valid for all* θ.

Proof. We rely on a calculation that appears in [O], Proposition 1.

Use the inequality $\big($[O], (2.4)$\big)$

$$\Big|\frac{1-e^{i\theta}}{1-\rho e^{i\theta}}\Big| \leq \frac{2}{1+\rho}\ \text{ for }\ \theta\in[0,2\pi], 0<\rho<1. \tag{2.2}$$

From (2.2)

$$\begin{aligned} \log|1-e^{i\theta}| &\le \log|1-\rho e^{i\theta}| + \log\frac{2}{1+\rho} \\ &= -\sum_{j=1}^{\infty}\frac{\rho^j}{j}\cos j\theta + \log\frac{2}{1+\rho} \\ &\le -\sum_{j=1}^{J}\frac{\rho^j}{j}\cos j\theta + \frac{\rho^J}{J(1-\rho)} + C(1-\rho) \end{aligned} \tag{2.3}$$

by partial summation and since

$$\log\frac{2}{1+\rho} = -\log\Big(1-\frac{1-\rho}{2}\Big).$$

Thus (2.1) follows from (2.3) with ρ as above.

□

Proposition 2.2. *There is a subset* $\{a_1 \dots a_m\} \subset \{1, \dots, n\}$ *of size*

$$m \asymp \frac{n}{2}$$

and

$$\Big\| \prod_{k=1}^{m} |1-z^{a_k}| \Big\|_{L^\infty(|z|=1)} \le e^{c\sqrt{n}\,\sqrt{\log n})(\log\log n)}. \tag{2.4}$$

Remark. (2.4) is a slight improvement of the estimate

$$\Big\| \prod_{k=1}^{m} |1-z^{a_k}| \Big\|_{L^\infty(|z|=1)} \le e^{c\sqrt{n}\log n}$$

resulting from a construction in [Kol1], p. 162 of a set $\{a_1, \dots, a_m\}$ as above and such that

$$\sum_{k=1}^{m}\cos 2\pi a_k\theta \ge -c\sqrt{m}$$

and Lemma 2.1

$$\begin{aligned} \log\prod_{k=1}^{m}|(1-2a_k)| &\le -\sum_{j=1}^{J}\frac{\rho^j}{j}\sum_{k=1}^{m}\cos 2\pi a_k(j\theta) + O\Big(\frac{m}{\sqrt{J}}\Big) \\ &\le C(logJ)\sqrt{m} + O\Big(\frac{m}{\sqrt{J}}\Big) \\ &< C\log n\,\sqrt{n}, \end{aligned}$$

taking $J = m^2$.

Proof of Proposition 2.2. Take independent selectors $(\xi_j)_{1\le j<n}$ with values $0, 1$ and mean $\mathbb{E}[\xi_j] = 1 - \frac{j}{n}$. Let $F_n(\theta) = 2\sum_{0<j<n}(1-\frac{j}{n})\cos 2\pi j\theta + 1$ be the Fejer kernel

$$\sum_{k=1}^m \cos a_k\theta = \sum_{\ell=1}^n \xi_\ell \cos \ell\theta = \frac{1}{2}F_n(\theta) - \frac{1}{2} + \sum_{\ell=1}^n (\xi_\ell - \mathbb{E}[\xi_\ell])\cos \ell\theta. \tag{2.5}$$

By Lemma 2.1 (applies with $J = n^{10}$)

$$\sum_{k=1}^m \log|1 - e^{2\pi i a_k\theta}| \le -\sum_{j=1}^J\sum_{k=1}^m \frac{\rho^j}{j}\cos 2\pi j a_k\theta + O\Big(\frac{m}{\sqrt{J}}\Big) \tag{2.6}$$

and we take J at least n to bound the last term in the right hand side of (2.5) by $\sqrt{n}$. We analyze the first term. Inserting (2.5) gives the sum of the following two expressions ((2.7) and (2.8))

$$-\sum_{j=1}^J \frac{\rho^j}{j}\Big(\frac{1}{2}F_n(j\theta) - \frac{1}{2}\Big) \tag{2.7}$$

$$-\sum_{j=1}^J\sum_{\ell=1}^n \frac{\rho^j}{j}(\xi_\ell - \mathbb{E}[\xi_\ell])\cos 2\pi\ell j\theta. \tag{2.8}$$

Since $F_n(j\theta) \ge 0$, (2.7) $\le \log J$.

Rewrite

$$(2.8) = -\sum_{\ell=1}^n (\xi_\ell - \mathbb{E}[\xi_\ell])\Big[\sum_{j=1}^J \frac{\rho^j}{j}\cos 2\pi j\ell\theta\Big]. \tag{2.9}$$

Note that all frequencies in (2.9) are bounded by nJ.

Applying the probabilistic Salem-Zygmund inequality [Kol3] shows that with large probability

$$(2.9) \lesssim \sqrt{\log\, nJ}\Big[\sum_{\ell=1}^n\Big|\sum_{j=1}^J \frac{\rho^j}{j}\cos 2\pi j\ell\theta\Big|^2\Big]^{\frac{1}{2}}. \tag{2.10}$$

Our next task is to evaluate the expression $\sum_{\ell=1}^n\Big|\sum_{j=1}^J \frac{\rho^j}{j}\cos 2\pi j\ell\theta\Big|^2$.

A first observation is that we can assume

$$\|\theta\| > \frac{1}{10n} \tag{2.11}$$

since otherwise

$$|1-e^{2\pi i a_k\theta}|\le 2\pi a_k\|\theta\|<\frac{2\pi}{10}<1$$

for all $k=1,\dots,m$, and also the left hand side of (2.4) is bounded by 1.

Next, we note that (since $\rho=1-\frac{1}{\sqrt{J}}$)

$$\begin{aligned}\Big|\sum_{j=1}^{J}\frac{\rho^j}{j}\cos 2\pi j\ell\theta\Big| &\le \big|\log|1-\rho e(\ell\theta)|\big|+\frac{\rho^J}{J(1-\rho)}\\ &< \big|\log|1-\rho e(\ell\theta)|\big|+1.\end{aligned}$$

Hence

$$\sum_{\ell=1}^{n}\Big|\sum_{j=1}^{J}\frac{\rho^j}{j}\cos 2\pi j\ell\theta\Big|^2 \lesssim \sum_{\ell=1}^{n}\big|\log|1-\rho e(\ell\theta)|\big|^2+n. \tag{2.12}$$

Fix θ and for $1<R\lesssim\log J$ define the dyadic set

$$S_R=\{1\le\ell\le n:\big|\log|1-\rho e(\ell\theta)|\big|\sim R\}.$$

Thus for $\ell\in S_R$

$$\|\ell\theta\|<|1-\rho e(\ell\theta)|<e^{-cR}=:\varepsilon.$$

Let $q\in\mathbb{N}$ be the smallest integer with $\|q\theta\|<2\varepsilon$. It follows that $|S_R|\lesssim\frac{n}{q}+1$. Assuming $q>R^3$, one obtains

$$\sum_{\ell\in S_R}\big|\log|1-\rho e(\ell\theta)|\big|^2\lesssim\Big(\frac{n}{R^3}+1\Big)R^2$$

with collected contribution (summing over dyadic R)

$$\sim n+(\log J)^2. \tag{2.13}$$

It remains to consider θ's with the property that for some large R and $q<R^3$,

$$\|q\theta\|<e^{-cR}.$$

Hence either θ admits a rational approximation

$$\Big|\theta-\frac{a}{q}\Big|<\frac{e^{-cR}}{q}<e^{-cR},\ \ q<R^3 \ \text{ and } \ (a,q)=1 \tag{2.14}$$

or $\big($in (2.14) when $a=0\big)$, by (2.11)

$$\frac{1}{n}\lesssim\|\theta\|<e^{-cR}. \tag{2.15}$$

Consider first the case (2.15). Then

$$|S_R| \leq |\{\ell = 1, \dots, n : \|\ell\theta\| < e^{-cR}\}| \lesssim ne^{-cR}$$

and the above estimate still holds.

Assume next that θ satisfies (2.14). Write

$$\theta = \frac{a}{q} + \psi \ \text{ with } \ \beta = |\psi| < e^{-cR}. \tag{2.16}$$

First, we consider the case $\beta \gtrsim \frac{1}{nq}$.

Let $V \subset \{1, \dots, n\}$ be an interval of size $\sim \frac{1}{q\beta}$ so that $\{\ell\theta : \ell \in V\}$ consists of $q\beta$-separated points filling a fraction of $[0, 1]$ (mod 1). Hence

$$\sum_{\ell \in V} \big| \log |1 - \rho e(\ell\theta)| \big|^2 \lesssim \frac{1}{\beta q} \int_0^1 \big| \log |1 - \rho e(t)| \big|^2 dt + \log^2(1 - \rho)$$
$$\lesssim \frac{1}{\beta q} + \log^2 J$$

and

$$\sum_{\ell=1}^{n} \big| \log |1 - \rho e(\ell\theta)| \big|^2 \lesssim n + nq\,\beta \log^2 n \lesssim n$$

unless

$$q\beta \log^2 n > 1, \text{ i.e. } \log n > e^{cR} \text{ or } R \lesssim \log\log n$$

where we used (2.14). Thus if $\beta \gtrsim \frac{1}{nq}$, (2.12) $\lesssim n(\log\log n)^2$.

The next case is $\beta < \frac{1}{100nq}$.

It follows that for $1 \leq \ell \leq n$

$$\Big| \ell\theta - \frac{\ell a}{q} \Big| < \frac{1}{100q}. \tag{2.17}$$

We obtain

$$\sum_{q \nmid \ell} \big| \log |1 - \rho e(\ell\theta)| \big|^2 \lesssim n \int_0^1 \big| \log |1 - \rho e(t)| \big|^2 dt \lesssim n$$

and

$$\begin{aligned}\sum_{q|\ell}\big|\log|1-\rho e(\ell\theta)|\big|^2 &\sim \frac{1}{q\beta}\int_0^{n\beta}\big|\log|1-\rho e(t)|\big|^2 dt\\ &\leq \frac{1}{q\beta}\int_0^{n\beta}\Big(\log\frac{1}{t}\Big)^2 dt \\ &\lesssim \frac{n}{q}(\log n\beta)^2.\end{aligned}\tag{2.18}$$

We obtain again a bound $O(n)$ unless

$$|\log n\beta| > \sqrt{q}$$

i.e.

$$\beta < \frac{e^{-\sqrt{q}}}{n}.\tag{2.19}$$

Thus (2.17) may be replaced by

$$\Big|\ell\theta - \ell\frac{a}{q}\Big| < e^{-\sqrt{q}} \text{ for } 1 \leq \ell \leq n.\tag{2.20}$$

For θ satisfying (2.20) we proceed in a different way. Write

$$\begin{aligned}\prod|1-e(a_k\theta)| &= \prod_{j=1}^n |1-e(j\theta)|^{\xi_j}\\ &\lesssim \prod_{j=1}^n\Big(\Big|1-e\Big(j\frac{a}{q}\Big)\Big| + \frac{1}{q^{10}}\Big)^{\xi_j}.\end{aligned}\tag{2.21}$$

We replace ξ_j by its expectation $\mathbb{E}[\xi_j] = 1-\frac{j}{n}$ using again a random argument. Thus if

$$\prod_{j=1}^n\Big(\Big|1-e\Big(j\frac{a}{q}\Big)\Big| + \frac{1}{q^{10}}\Big)^{1-\frac{j}{n}}\tag{2.22}$$

we have

$$|\log(2.21) - \log(2.22)| \leq \Big|\sum_{j=1}^n\Big(\xi_j - \mathbb{E}[\xi_j]\Big)\log\Big(\Big|1-e\Big(j\frac{a}{q}\Big)\Big| + \frac{1}{q^{10}}\Big)\Big|.\tag{2.23}$$

Recall that $q < R^3 \lesssim (\log J)^3 \sim (\log n)^3$. Thus with high probability we may bound (2.23) by $c\sqrt{n}\,\sqrt{\log\log n}\log q < c\sqrt{n}(\log\log n)^3$.

Hence

$$(2.21) \ \leq \ e^{c\sqrt{n}(\log\log n)^3}(2.22).$$

Partition $\{1, \ldots, n\}$ in intervals $I = [rq, (r+1)q - 1]$ and estimate for each such interval

$$\begin{aligned}
&\prod_{j\in I}\Big(\Big|1 - e\Big(j\frac{a}{q}\Big)\Big| + \frac{1}{q^{10}}\Big)^{1-\frac{j}{n}}\\
&\leq q^{c\frac{q^2}{n}}\Big[\frac{1}{q^{10}}\prod_{s=1}^{q-1}\Big(\Big|1 - e\Big(s\frac{a}{q}\Big)\Big| + \frac{1}{q^{10}}\Big]^{1-\frac{rq}{n}} \qquad (2.24)\\
&\leq q^{c\frac{q^2}{n}}\Big[\frac{1}{q^{10}}\prod_{s=1}^{q-1}\Big|1 - e\Big(\frac{s}{q}\Big)\Big|\Big]^{1-\frac{rq}{n}}.
\end{aligned}$$

The product $\prod_{s=1}^{q-1}\big|1 - e\big(\frac{s}{q}\big)\big|$ may be evaluated using Lemma 2.1 taking $J = q^2, \rho = 1 - \frac{1}{q}$. Thus clearly

$$\begin{aligned}
\sum_{s=1}^{q-1}\log\Big|1 - e\Big(\frac{s}{q}\Big)\Big| &\leq -\sum_{j=1}^{J}\frac{\rho^j}{j}\sum_{s=1}^{q-1}\cos 2\pi j\frac{s}{q} + O(1)\\
&\leq \sum_{\substack{1\leq j\leq J\\ q\nmid j}}\frac{\rho^j}{j} + q\sum_{\substack{1\leq j\leq J\\ q\mid j}}\frac{\rho^j}{j} + O(1)\\
&< \log q + C
\end{aligned}$$

implying that

$$(2.24) < q^{c\frac{q^2}{n}}\Big(\frac{1}{q^{10}}e^{\log q + c}\Big)^{1-\frac{rq}{n}} < q^{c\frac{q^2}{n}}. \qquad (2.25)$$

Since (2.22) is obtained as product of (2.24), (2.25) over the intervals I, we showed that

$$(2.22) < q^{c\frac{q^2}{n}n}2^q < e^{(\log n)^3}.$$

Thus the preceding shows that if θ satisfies (2.20), then

$$\prod|1 - e(a_k\theta)| < e^{c\sqrt{n}(\log\log n)^3}. \qquad (2.26)$$

Going back to (2.10), omitting the case (2.20) estimated by (2.26), we obtained the bound $cn(\log\log n)^2$ on (2.12) which permits to majorize (2.8) by $c\sqrt{n\log n}(\log\log n)$ and $\prod|1 - e(a_k\theta)|$ by $e^{c\sqrt{n\log n}\log\log n}$. This completes the proof of Proposition 2.2.

3. Almost full proportion

It was observed in [E-S] that

$$\lim_{n\to\infty} M(1,\dots,n)^{\frac{1}{n}} \tag{3.1}$$

exists and lies strictly between 1 and 2.

This fact is in contrast with Proposition 2.2 which gives a subset $S\subset\{1,\dots,n\}, |S|\asymp\frac{n}{2}$ s.t.

$$\log M(S)\lesssim\sqrt{n}(\log n)^{\frac{1}{2}}\log\log n. \tag{3.2}$$

However

Proposition 3.1. *There is a constant $\tau>0$ such that if $S\subset\{1,\dots,n\}$ satisfies $|S|>(1-\tau)n$, then*

$$\log M(S)>cn \tag{3.3}$$

for some $c>0$.

Thus (3.3) generalizes (3.1) in some sense, but in view of (3.2), it fails dramatically if we do not assume $1-\frac{|S|}{n}$ small enough.

Proof of Proposition 3.1.

It will be convenient to use Fact 2 for an appropriate μ-convolution, which allow us to estimate the tail contribution in the k-summation.

Thus consider

$$-\min_\theta\Big\{\sum_{j\in S}\sum_{k=1}^\infty\frac{\cos 2\pi kj\cdot}{k}*\mu\Big\}(\theta)$$

$$=-\min_\theta\sum_{k=1}^\infty\sum_{j\in S}\frac{\hat\mu(jk)}{k}\cos 2\pi kj\theta$$

$$\geq-\min_\theta\sum_{k=1}^{k_0}\sum_{j=1}^{n}\frac{\hat\mu(jk)}{k}\cos 2\pi kj\theta \tag{3.4}$$

$$-(\log k_0)\pi n$$

$$-\sum_{k>k_0}\sum_{j=1}^{n}\frac{|\hat\mu(jk)|}{k} \tag{3.5}$$

since we assumed $|S| > (1-\tau)n$.

Separating in (3.4) the cases $k=1$, and $2 \le k \le k_0$, we write

$$\begin{aligned}(3.4) \ge &-\Big(\sum_{j=1}^n \cos 2\pi j\theta\Big) - \sum_{j=1}^n |1-\hat\mu(j)|\\ &-\sum_{k=2}^{k_0} \frac{1}{k}\Big|\sum_{j=1}^n \hat\mu(jk)\cos 2\pi kj\theta\Big|.\end{aligned} \tag{3.6}$$

Take $\mu = F_{nR}(\theta)$, $R>1$ an appropriate constant and $F_{nR}(\theta)$ the Féjer kernel.

Thus

$$\begin{aligned}\widehat{F}_{nR}(s) &= 1 - \frac{|s|}{nR} \quad \text{for } \ |s| \le nR\\ &= 0 \qquad\qquad \text{otherwise.}\end{aligned}$$

Take $\theta = \frac{3}{4n}$. The first term in (3.6) becomes, since

$$\sum_{j=1}^n \cos jx = \frac{1}{2}D_n(x) - \frac{1}{2}, \quad \text{where } \ D_n(x) = \frac{\sin(n+\frac12)x}{\sin\frac{x}{2}}$$

is the Dirichlet kernel,

$$\frac12 - \frac12 \frac{\sin\frac{3\pi}{2n}(n+\frac12)}{\sin\frac{3\pi}{4n}} \sim +\frac{1}{2\sin\frac{3\pi}{4n}}.$$

The second term is

$$-\sum_{j=1}^n \frac{j}{nR} = -\frac{n+1}{2R}.$$

The third term becomes

$$-\sum_{k=2}^{k_0}\frac1k\Big|\sum_{j=1}^n\Big(1-\frac{jk}{nR}\Big)_+ \cos\pi\frac{3kj}{2n}\Big|. \tag{3.7}$$

By partial summation, the inner sum is bounded by

$$\begin{aligned}&\max_{j_1\le \min(n,\frac{nR}{k})}\Big|\sum_{j=1}^{j_1}\cos\pi\frac{3kj}{2n}\Big|\\ &= \max_{j_1\le \min(n,\frac{nR}{k})}\Big|\frac12 D_{j_1}\Big(\frac32\pi\frac kn\Big) - \frac12\Big|\\ &\le \frac{1}{2|\sin\frac34\pi\frac kn|} + \frac12.\end{aligned}$$

For $k < k_0 = o(n)$, the first term

$$\sim \frac{1}{2k \sin \frac{3\pi}{4n}}.$$

Hence

$$\begin{aligned}(3.7) &\geq -\sum_{k=2}^{k_0} \frac{1}{2k^2} \frac{1}{\sin \frac{3\pi}{4n}} - \log k_0 \\ &\geq -\frac{1}{2 \sin \frac{3\pi}{4n}} \Big(\frac{\pi^2}{6} - 1\Big) - \log k_0.\end{aligned}$$

It follows from the preceding that

$$\begin{aligned}(3.4) &\geq +\frac{1}{2 \sin \frac{3\pi}{4n}} \Big(2 - \frac{\pi^2}{6}\Big) - \log k_0 - \frac{n+1}{2R} \\ &= cn - \log k_0\end{aligned}$$

for R a sufficiently large constant.

We bound (3.5) by

$$(3.5) \geq -\sum_{k \geq k_0} \frac{1}{k} \sum_{j \leq \frac{nR}{k}} 1 \geq -\sum_{k \geq k_0} \frac{nR}{k^2} \geq -\frac{R}{k_0} n.$$

In summary, we proved that

$$-\sum_{j \in S} \sum_{k=1}^{\infty} \frac{\hat{\mu}(jk)}{k} \cos 2\pi jk \, \frac{3}{4n} \geq cn - \log k_0 - \tau(\log k_0)n - \frac{C'n}{k_0} > \frac{c}{2} n$$

be choosing first k_0 large enough and then assuming τ sufficiently small.

This proves Proposition 3.1.

4. Sets with large arithmetical Diameter

As we pointed out the general lower bound $M(a_1, \dots, a_n) > \sqrt{n}$ remains unimproved. However Proposition 4.1 stated below shows that in certain cases one can do better.

First, we give the following definition.

Definition. $D = \{v_1, \dots, v_m\} \subset \mathbb{Z}$ *is called dissociated provided the relation*

$$\varepsilon_1 v_1 + \cdots + \varepsilon_m v_m = 0 \quad \textit{with } \varepsilon_i = 0, 1, -1$$

implies that $\varepsilon_1 = \cdots = \varepsilon_m = 0$.

We note that Hadamard lacunary sets are dissociated.

Proposition 4.1. *Assume $S = \{a, \dots, a_n\}$ contains a dissociated set D of size m. Then*

$$\log M(a_1, \dots, a_n) \gg \frac{m^{\frac{1}{2}-o(1)}}{(\log n)^{\frac{1}{2}}}. \tag{4.1}$$

Thus (4.1) improves the general lower bound from [E-S] provided $m > (\log n)^{3+\varepsilon}$.

Remark. By a result of Pisier [P], our assumption is equivalent to S containing a Sidon set Λ of size $|\Lambda| \sim m$. Here 'Sidon set' is in the harmonic analysis sense i.e.

$$\Big\| \sum_{n\in\Lambda} \lambda_n e(n\theta) \Big\|_\infty \geq c \sum |\lambda_n| \text{ for all scalars } \{\lambda_n\}$$

with $c = c(\Lambda)$ to be considered as a constant. (This concept is different from the Sidon sets in combinatorics!).

Dissociated sets are Sidon and conversely, Pisier proved that if Λ is a finite Sidon set, then Λ contains a proportional dissociated set.

Proof of Proposition 4.1.

We derive (4.1) from the equivalent statement

$$\max_\theta \big(\log|1 - e(a_1\theta)| + \cdots + \log|1 - e(a_n\theta)|\big) \gg \frac{m^{\frac{1}{2}-o(1)}}{(\log n)^{1/2}} \tag{4.2}$$

which, since $\int \log|1 - e(a\theta)| = 0$ for $a \in \mathbb{Z}\backslash\{0\}$, is a consequence of the stronger claim that

$$\|F\|_1 \gg \frac{m^{\frac{1}{2}-o(1)}}{(\log n)^{1/2}} \tag{4.3}$$

denoting

$$F(\theta) = \log|1 - e(a_1\theta)| + \cdots + \log|1 - e(a_n\theta)|.$$

Recall that by Fact 1

$$F(\theta) = -\sum_{k=1}^{\infty} \frac{1}{k} f(k\theta) \tag{4.4}$$

with

$$f(\theta) = \sum_{j=1}^{n} \cos(2\pi a_j\theta).$$

We first perform a finite Mobius inversion on (4.4). Recall that

$$\sum_{\substack{d|k, d\leq r\\ d \text{ square free}}} \mu(d) = \begin{cases} 1 & \text{if } k=1 \\ 0 & \text{if } 1<k\leq r \end{cases}$$

Hence

$$\begin{aligned} \sum_{\substack{d<r\\ \text{square free}}} F(d\theta)\frac{\mu(d)}{d} &= -\sum_{j=1}^{n}\sum_{k=1}^{\infty}\sum_{\substack{d<r\\ \text{square free}}} \cos(2\pi a_j dk\theta)\frac{\mu(d)}{dk} \\ &= -\sum_{j=1}^{n}\sum_{\ell=1}^{\infty}\frac{\cos(2\pi a_j \ell\theta)}{\ell}\Big[\sum_{\substack{d|\ell, d<r\\ \text{square free}}}\mu(d)\Big] \\ &= -f(\theta) - \sum_{j=1}^{n}\sum_{\ell>r}\frac{\cos(2\pi a_j \ell\theta)}{\ell}\Big[\sum_{\substack{d|\ell, d<r\\ \text{square free}}}\mu(d)\Big] \\ &= -f(\theta)+G(\theta), \end{aligned} \tag{4.5}$$

where

$$G(\theta) = -\sum_{j=1}^{n}\sum_{\ell>r}\frac{\cos(2\pi a_j \ell\theta)}{\ell}\Big[\sum_{\substack{d|\ell, d<r\\ \text{square free}}}\mu(d)\Big].$$

Note also that

$$\Big|\sum_{\substack{d|\ell, d<r\\ \text{square free}}}\mu(d)\Big| \leq 2^{\omega(\ell)}, \tag{4.6}$$

where $\omega(\ell)$ is the number of distinct prime factors of ℓ.

Denote m the size of the largest dissociated set contained in $\{a_1, \dots, a_n\}$. Our first task will be to bound the Fourier transform $\|\hat{G}\|_\infty$ of G.

Thus given $t \in \mathbb{Z}$, we have

$$|\hat{G}(t)| \leq \frac{1}{2}\sum_{j=1}^{n}\frac{a_j}{t}2^{\omega(\frac{t}{a_j})}. \tag{4.7}$$

We will bound (4.7) by considering dyadic ranges, letting for $K>r$ dyadic

$$J = J_K = \{j\in[1,n] \ : \ a_j|t \quad \text{and} \quad \frac{t}{a_j}\sim K\}.$$

Thus

$$\sum_{j\in J}\frac{a_j}{t}2^{\omega(\frac{t}{a_j})}\leq\sqrt{\sum_{j\in J}\Big(\frac{a_j}{t}\Big)^2}\Big(\sum_{k\leq K}4^{\omega(k)}\Big)^{\frac12}$$
$$\lesssim|J|^{\frac12}K^{-1}K^{\frac12}(\log K)^2=\Big(\frac{|J|}{K}\Big)^{\frac12}(\log K)^2. \tag{4.8}$$

Assume

$$|J|>\frac{K}{(\log K)^8}. \tag{4.9}$$

Our aim is to get a contradiction for appropriate choice of r.

At this point, we invoke the following result from [H-T] (see Fq (1.14)).

Denote

$$\psi(x,y)=\big|\{n\leq x:\ \text{if}\ \ p|n,\ \text{then}\ p\leq y\}\big|.$$

Lemma 4.2. *For any $0<\alpha<1$, we have*

$$\psi\big(x,(\log x)^{1/\alpha}\big)<x^{1-\alpha+o(1)}\ \textit{for}\ \ x\to\infty. \tag{4.10}$$

It follows from (4.9) that for any fixed $1>\alpha>0$, we have

$$|J|>2\psi\big(K,(\log K)^{\frac1\alpha}\big). \tag{4.11}$$

We make the following construction.

By (4.11), there is $j_1\in J$ such that $\frac{t}{a_{j_1}}$ has a prime divisor $p_1>(\log K)^{\frac1\alpha}$ and we write $\frac{t}{a_{j_1}}=p_1b_1$.

Next, let $J_1=\{j\in J:p_1|\frac{t}{a_j}\ \}$. Hence $|J_1|<\frac{K}{p_1}+1<\frac{K}{(\log K)^{\frac1\alpha}}<\frac{|J|}{(\log K)^{\frac1\alpha-8}}$ where we assume α taken much smaller than $\frac18$.

It follows that also

$$|J\backslash J_1|>\Big(2-\frac{1}{(\log K)^{\frac1\alpha-8}}\Big)\psi\big(K,(\log K)^{\frac1\alpha}\big)$$

which permits to introduce $j_2\in J\backslash J_1$ and a prime $p_2>(\log K)^{\frac1\alpha}$ such that $p_2|\frac{t}{a_{j_2}}$. Write $\frac{t}{a_{j_2}}=p_2b_2$. Clearly $p_2\neq p_1$ and $p_1\nmid b_2$.

The contribution of the process is clear. We may introduce elements

$$j_1,\dots,j_s\in J\ \text{ with }\ s\gtrsim(\log K)^{\frac1\alpha-8}$$

and prime divisors $p_{s'}|\frac{t}{a_{j_{s'}}}$. Write $\frac{t}{a_{j_{s'}}} = p_{s'}b_{s'}$ such that $p_{s'} \nmid \frac{t}{a_{j_{s''}}}$ for $s' < s''$. Hence $p_{s''} \neq p_{s'}$ for $s' \neq s''$ and

$$p_{s'} \nmid b_{s''} \text{ for } s' < s''. \tag{4.12}$$

We claim that the set $\{a_{j_1}, \ldots, a_{j_s}\}$ is dissociated. Otherwise, there is a non-trivial relation

$$\varepsilon_1 a_{j_1} + \cdots + \varepsilon_s a_{j_s} = 0 \qquad \text{with } \varepsilon_{s'} = 0, 1, -1$$

which by the preceding translates in

$$\varepsilon_1 \frac{1}{p_1 b_1} + \cdots + \varepsilon_s \frac{1}{p_s b_s} = 0$$

or

$$\sum_{s'=1}^{s} \varepsilon_{s'} \prod_{s'' \neq s'} p_{s''} b_{s''} = 0.$$

Let s_1 be the smallest s' with $\varepsilon_{s'} \neq 0$. Then

$$\sum_{s'=s_1}^{s} \varepsilon_{s'} \prod_{\substack{s'' \neq s' \\ s'' \geq s_1}} p_{s''} b_{s''} = 0. \tag{4.13}$$

Since

$$p_{s_1} \Big| \prod_{\substack{s'' \neq s' \\ s'' \geq s_1}} p_{s''} b_{s''} \text{ for } s' > s_1,$$

identity (4.13) implies

$$p_{s_1} \Big| \prod_{s'' > s_1} b_{s''},$$

contradicting (4.12).

Hence $\{a_{j_1}, \ldots, a_{j_s}\}$ is dissociated and by definition of m,

$$s \leq m$$

implying

$$m \geq (\log K)^{\frac{1}{\alpha} - 8} \quad \text{and} \quad \log r \leq m^{\frac{\alpha}{1-8\alpha}}.$$

Thus, by taking

$$\log r \sim m^{2\alpha} \quad (\alpha \text{ small enough})$$

we obtain a contradiction under assumption (4.9).

Hence

$$|J_K| < \frac{K}{(\log K)^8} \quad \text{for } K > r$$

and summing (4.8) over dyadic ranges of $K > r$ gives the bound

$$|\hat{G}(t)| < \sum_{\substack{K>r \\ \text{dyadic}}} \frac{1}{(\log K)^2} \lesssim \frac{1}{\log r}. \tag{4.14}$$

Consequently

$$\widehat{(4.5)}(t) = -\hat{f}(t) + O\Big(\frac{1}{\log r}\Big) = -\hat{f}(t) + o(1) \text{ for all } t \in \mathbb{Z}. \tag{4.15}$$

Since

$$\hat{f}(j) = \frac{1}{2},$$

we have

$$\widehat{(4.5)}(j) = -\frac{1}{2} + o(1). \tag{4.16}$$

Next, let D be a size m dissociated set in $\{a_1, \dots, a_n\}$. Define

$$\varphi(\theta) = \frac{1}{\sqrt{m}} \sum_{j \in D} e(j\theta).$$

Also, let Φ, Ψ be the dual Orliez functions

$$\Phi(x) = x\sqrt{\log(2+x)} \quad \text{and} \quad \Psi(x) = e^{x^2}.$$

It is well known (e.g. Theorem 3.1 in [Rud].) that

$$\|\varphi\|_{L^\Psi} = O(1).$$

By (4.16)

$$\Big(\frac{1}{2} - o(1)\Big)\sqrt{m} \le \Big| \int_0^1 (4.5)\varphi(\theta) d\theta \Big| \le C\|(4.5)\|_{L^\Phi} \tag{4.17}$$

It remains to bound $\|(4.5)\|_{L^\Phi}$.

Estimate

$$\begin{aligned} &\int |(4.5)| \sqrt{\log(|(4.5)| + 2)}\, d\theta \\ \le\ &\sum_{j>0} 2^{j/2} \int_{2^{2^{j-1}} \le \lambda \le 2^{2^j}} \mu(M)\, d\lambda, \end{aligned} \tag{4.18}$$

Where $M = \{\theta : (4.5)(\theta) > \lambda\}$ and μ is the measure. Using the left hand side of (4.5), the j-summands is bounded by

$$2^{j/2}\|(4.5)\|_1 \lesssim 2^{j/2}\log r\ \|F\|_1. \tag{4.19}$$

Also, let $\Psi_1(u) = e^u$. Then

$$\Big\|\sum_{d\leq r}\frac{|F(d\theta)|}{d}\Big\|_{L^{\Psi_1}} \leq (\log r)\|F\|_{L^{\Psi_1}} \lesssim n\log r,$$

since $\|\log|1-e^{i\theta}|\,\|_{L^{\Psi_1}} < \infty$.

Thus also the bound

$$\mu(M) \leq e^{-c\,\frac{\lambda}{n\log r}}$$

implying the following bound for the j-summands

$$2^{j/2}2^{2^j}e^{-c\,\frac{2^{2^{j-1}}}{n\log r}}. \tag{4.20}$$

Hence

$$(4.18) < \sum_j 2^{j/2}\min\Big((\log r)\|F\|_1, 2^{2^j}e^{-c\,\frac{2^{2^{j-1}}}{n\log r}}\Big).$$

For $2^{2^{j-2}} < n\log r$, we get the contribution

$$(\log n)^{\frac{1}{2}}\log r\|F\|_1.$$

For $2^{2^{j-2}} \geq n\log r$, we bound by

$$(n\log r)^{4+\epsilon}e^{-cn\log r} + (n\log r)^{4\cdot 2+\epsilon}\,e^{-c(n(\log r))^3} + \cdots + (n\log r)^{4\cdot 2^{u-1}+\epsilon}\,e^{-c(n\log r)^{2^u-1}} + \cdots$$
$$< O(1).$$

Hence

$$\|(4.5)\|_{L^\Phi} \lesssim (4.18) < (\log n)^{\frac{1}{2}}m^{2\alpha}\|F\|_1 \tag{4.21}$$

recalling above choice for $\log r$.

Returning to (4.17), we proved that

$$\Big(\frac{1}{2} - o(1)\Big)m^{\frac{1}{2}-2\alpha} \lesssim (\log n)^{\frac{1}{2}}\|F\|_1$$

hence

$$\|F\|_1 \gtrsim m^{\frac{1}{2}-\varepsilon}(\log n)^{-\frac{1}{2}}.$$

This proves (4.3) and hence Proposition 4.1.

Acknowledgment

During the preparation of this paper, the first author was partially supported by the NSF Grant DMS 1301619 and the second author by the NSF Grant DMS 1301608.

REFERENCES

[A] F.V. Atkinson, *On a problem of Erdos and Szekeres*, Canad. Math. Bull., 4 (1961), 7–12.

[C1] S. Chowla, *The Riemann zeta and allied functions*, Bull. Amer. Math. Soc, 58 (1952), 287–305.

[C2] S. Chowla, *Some applications of a method of A. Selberg*, J. reine angew. Math., 217 (1965), 128–132.

[E-S] P. Erdos, G. Szekeres, *On the product* $\prod_{k=1}^{n}(1-z^{a_k})$, Publ. de l'Institut mathematique, 1950.

[K-O] Y. Katznelson, D. Ornstein, *The differentiability of the conjugation of certain diffeomorphisms to the circle*, ETDA 9 (1989), no 4, 643–680.

[Kol1] M.N. Kolountzakis, *On nonnegative cosine polynomials with nonnegative integral coefficients*, Proc. AMS 120 (1994), 157–163.

[Kol2] M.N. Kolountzakis, *A construction related to the cosine problem*, Proc. Amer. Math. Soc. 122 (1994), vol. 4, 1115–1119.

[Kol3] M.N. Kolountzakis, *Some applications of probability to additive number theory and harmonic analysis*, Number theory (New York Seminar, 1991-1995),Springer, New York, (1996), 229–251.

[Kol4] M.N. Kolountzakis, *The density of* $B_h[g]$ *sets and the minimum of dense cosine sums*, J. Number Theory 56 (1996), 1, 4-11.

[H-T] A. Hildebrand, G. Tenenbaum, *Integers without large prime factors*, J. Theorie des Nombres de Bordeaux, 5 (1993), no 2, 411–484.

[O] A.M. Odlyzko, *Minima of cosine sums and maxima of polynomials on the unit circle*, J. London Math. Soc (2), 26 (1982), no 3, 412–420.

[P] G. Pisier, *Arithmetic Characterization of Sidon Sets*, 8 (1983), Bull. AMS, 87-89.

[Rud] W. Rudin, *Trigonometric series with gaps*, J. Math. Mech., 9 (1960), 203–227.

[R] I.Z. Rusza, *Negative values of cosine sums*, Acta Arith. 111 (2004), 179-186.

[S] A. Schinzel, *On the number of irreducible factors of a polynomial*, Topics in number theory (ed. P. Turan, North Holland, Amsterdam, (1976), 305–314.

INSTITUTE FOR ADVANCED STUDY, PRINCETON, NJ 08540

E-mail address: bourgain@math.ias.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, RIVERSIDE, CA 92507

E-mail address: mcc@math.ucr.edu