SUMS AND PRODUCTS FROM A FINITE SET OF REAL NUMBERS

Kevin Ford

Dedicated to the memory of Paul Erdös

Abstract. If A is a finite set of positive integers, let $E_h(A)$ denote the set of h-fold sums and h-fold products of elements of A. This paper is concerned with the behavior of the function $f_h(k)$, the minimum of $|E_h(A)|$ taken over all A with $|A| = k$. Upper and lower bounds for $f_h(k)$ are proved, improving bounds given by Erdös, Szemerédi, and Nathanson. Moreover, the lower bound holds when we allow A to be a finite set of arbitrary positive real numbers.

For finite sets of real numbers A and B, define

$$A + B = \{a + b : a \in A, b \in B\}, \qquad AB = \{ab : a \in A, b \in B\}.$$

More generally, if $h \geqslant 2$ define

$$hA = \{a_1 + \cdots + a_h : a_i \in A\}, \qquad A^h = \{a_1 \cdots a_h : a_i \in A\}.$$

Erdös [E] conjectured that for any finite set A of positive integers,

$$|E_h(A)| \gg_\varepsilon |A|^{h-\varepsilon}, \tag{1}$$

where

$$E_h(A) = hA \cup A^h.$$

In other words, no set A can have simultaneously few sums and few products. Notice that trivially

$$\tfrac{1}{2}(|hA| + |A^h|) \leqslant |E_h(A)| \leqslant |hA| + |A^h|. \tag{2}$$

Our chief interest here is the behavior of the function

$$f_h(k) = \min\{|E_h(A)| : |A| = k, A \subset \mathbb{N}\}.$$

Erdös and Szemerédi [ES] proved the non-trivial bounds

$$k^{1+\delta} \ll f_2(k) \ll k^{2-c/\log_2 k}. \tag{3}$$

1991 *Mathematics Subject Classification.* Primary–11B75..
Key words and phrases. Sequences, sums and products..

Typeset by $\mathcal{AMS}$-TEX

where c and δ are positive constants an $\log_k x$ denotes the kth iterate of the logarithm. Nathanson [N] showed that $\delta = 1/31$ is admissible, and we note that the argument works for any finite set of positive *real* numbers. No bounds for $|E_h(A)|$ for $h \geqslant 3$ have been published. However, for any $a \in A$, A^h contains $a^{h-2}p$ for each $p \in A^2$ and hA contains $(h-2)a + s$ for each $s \in 2A$. Thus, by (2),

$$|E_h(A)| \geqslant \tfrac{1}{2}(|hA| + |A^h|) \geqslant \tfrac{1}{2}(|2A| + |A^2|) \geqslant \tfrac{1}{2}|E_2(A)|. \tag{4}$$

We also have

$$|E_h(A)| \leqslant |hA| + |A^h| \leqslant |A|^{h-2}(|2A| + |A^2|) \leqslant 2|A|^{h-2}|E_2(A)|.$$

In particular, if (1) fails for a particular h, it fails for all larger h.

When $h = 2$, (1) has been established for certain very special sets of positive integers A. Nathanson and Tenenbaum [NT] proved (1) under the assumption that $|2A| \leqslant 3|A| - 4$ using Freiman's structure theory of set addition (see [F]). As noted by Nathanson and Jia [NJ], (1) can also be proved in the case where A is contained in a "short" interval of length $|A|^{o(\log_2 |A|)}$ using the fact that $\log d(n) = O(\log n / \log_2 n)$, where $d(n)$ is the number of divisors of n.

In this note, we improve the lower bound for $|E_2(A)|$ using a refinement of Nathanson's argument [N].

Theorem 1. *If A is a finite set of positive real numbers, then*

$$|E_2(A)| \geqslant \tfrac{1}{6}|A|^{1+1/15}.$$

A slight modification of one part of the argument produces lower bounds for $|E_h(A)|$ for $h \geqslant 3$ which are superior to the bound obtained by combining (4) with Theorem 1. However, the exponent only tends to $8/7$ as h tends to infinity.

Theorem 2. *If A is a finite set of positive real numbers, then*

$$|E_h(A)| \gg |A|^{1+\frac{h-1}{7h+1}}.$$

Lastly, we investigate how small the sets $E_h(A)$ can be. Erdös and Szemerédi proved the lower bound in (3) by taking A to be a set of sufficiently "smooth" numbers (numbers without large prime factors). Using modern results concerning the distribution of smooth numbers, we prove an analogous result for $f_h(k)$, where the "constant" c grows rapidly with h.

Theorem 3. *For each fixed h, we have*

$$f_h(k) \leqslant k^{h - c_h/\log_2 k + O((\log_3 k)/(\log_2 k)^2)},$$

where $c_h = h(h-1)\log h$.

The starting point for the proof of Theorems 1 and 2 is a lower bound on the number of sums and products when B is contained in a dyadic interval. In this case, Nathanson [N] showed that $|E_2(B)| \gg |B|^{16/15}$.

Lemma 1. *Suppose B is a finite set of real numbers contained in $[x, 2x]$ for some positive x. Then*

$$|2B| + |B^2| \geqslant \tfrac{7}{20}|B|^{8/7}.$$

Proof. Let $k = |B|$ and suppose $k \geqslant 10^7$, for otherwise the right side in the lemma is less than $4k-2$ and the lemma is trivial. Suppose $1 \leqslant l < k$ and group the numbers in B as follows. Let B_1 be the set of l smallest numbers in B, let B_2 denote the set of l smallest numbers in $B \backslash B_1$, etc. This partitions B into $B_1, B_2, \dots, B_{[k/l]}$ with $< l$ numbers left over. Let the diameter of a set be the difference between the largest and the smallest numbers in the set. Let B^* be the set B_i with smallest diameter and let d be the diameter of B^*.

Now suppose $1 \leqslant i < j \leqslant [k/l]$ with $j - i \geqslant 3$ and

$$b_1^*, b_2^* \in B^*, \qquad b_i \in B_i, \qquad b_j \in B_j.$$

Then

$$b_1^* + b_i < (b_2^* + d) + (b_j - 2d) < b_2^* + b_j \tag{5}$$

and

$$\begin{aligned} b_j b_2^* &> (b_i + 2d) b_2^* \\ &\geqslant b_i(b_1^* - d) + 2d b_2^* \\ &= b_i b_1^* + d(2b_2^* - b_i) \geqslant b_i b_1^*. \end{aligned} \tag{6}$$

From now on consider only the sets $B_1, B_4, B_7, \dots$. By (5) and (6), the sets $B^* + B_i$ are distinct, as are the sets $B^* B_i$. Let

$$P_i = |B^* \cdot B_i|, \qquad S_i = |B^* + B_i|. \tag{7}$$

Then

$$|2B| + |B^2| \geqslant \sum_{i \equiv 1 \pmod 3} P_i + S_i. \tag{8}$$

Fix i and define

$$r(m) = |\{(b^*, b_i) : b^* b_i = m, b^* \in B^*, b_i \in B_i\}|.$$

When $r(m) > 0$, denote by $(b_j^*, b_j') \quad (1 \leqslant j \leqslant r(m))$ the distinct pairs of numbers $b_j^* \in B^*, b_j' \in B_i$ with product m. Notice that $b_{j_1}^* + b_{j_2}' \in B^* + B_i$ for each of the $r(m)^2$ pairs (j_1, j_2). For each $n \in B^* + B_i$, define

$$s_m(n) = |\{(j_1, j_2) : b_{j_1}^* + b_{j_2}' = n\}|.$$

With m, n fixed there are $\frac{1}{2}(s_m(n)^2 - s_m(n)) \geqslant s_m(n) - 1$ quadruples (j_1, j_2, j_3, j_4) with $b_{j_1}^* < b_{j_3}^*$ and

$$\begin{gathered} b_{j_1}^* + b_{j_2}' = b_{j_3}^* + b_{j_4}' = n, \\ b_{j_2}^* b_{j_2}' = b_{j_4}^* b_{j_4}' = m. \end{gathered} \tag{9}$$

On the other hand, given any four numbers $(b^*_{j_1}, b^*_{j_2}, b^*_{j_3}, b^*_{j_4})$ in B^* with $b^*_{j_1} < b^*_{j_3}$, equations (9) have at most one solution b'_{j_2}, b'_{j_4} and thus i, m and n are uniquely determined. If we let N_i be the number of quadruples corresponding to each i, then by (7) and the Cauchy-Schwarz inequality,

$$\begin{aligned} N_i &\geqslant \sum_m \sum_n s_m(n) - 1 \\ &\geqslant \sum_m (r(m)^2 - S_i) \\ &\geqslant l^4/P_i - P_i S_i. \end{aligned}$$

Also, $N_i \geqslant 0$ for each i. If $b^*_{j_1} < b^*_{j_3}$, then (9) implies $b^*_{j_2} < b^*_{j_4}$ and hence

$$\sum_i N_i \leqslant \tfrac{1}{4} l^4. \tag{10}$$

Define

$$\begin{aligned} I_1 &= \{ i \equiv 1 \pmod 3 : S_i P_i^2 \geqslant \tfrac{1}{2} l^4 \}, \\ I_2 &= \{ i \equiv 1 \pmod 3 : S_i P_i^2 < \tfrac{1}{2} l^4 \}. \end{aligned}$$

A straightforward calculation shows that

$$S_i + P_i \geqslant \frac{3}{2} l^{4/3} \qquad (i \in I_1). \tag{11}$$

We also have $N_i \geqslant l^4/2P_i$ for $i \in I_2$, hence by (10),

$$\sum_{i \in I_2} \frac{1}{P_i} \leqslant \frac{1}{2}. \tag{12}$$

Let $M_1 = |I_1|$, $M_2 = |I_2|$ and $H = M_1 + M_2$. By (8), (11), (12) and the Cauchy-Schwarz inequality,

$$\begin{aligned} |2B| + |B^2| &\geqslant \tfrac{3}{2} l^{4/3} M_1 + \sum_{i \in I_2} P_i \\ &\geqslant \tfrac{3}{2} M_1 l^{4/3} + 2M_2^2 \\ &= \tfrac{3}{2} l^{4/3} (H - M_2) + 2M_2^2. \end{aligned}$$

The right side is minimized at $M_2 = \frac{3}{8} l^{4/3}$. Since $H \geqslant \frac{1}{3}[k/l] \geqslant \frac{k}{3l} - \frac{1}{3}$, we obtain

$$\begin{aligned} |2B| + |B^2| &\geqslant \tfrac{3}{2} H l^{4/3} - \tfrac{9}{32} l^{8/3} \\ &\geqslant \tfrac{1}{2} k l^{1/3} - \tfrac{9}{32} l^{8/3} - \tfrac{1}{2} l^{4/3}. \end{aligned} \tag{13}$$

Ignoring the last term, the optimal value of l is

$$l = \left[\left(\tfrac{2}{9} k \right)^{3/7} \right].$$

The lemma now follows from (13), since $k \geqslant 10^7$ and $l \geqslant (\frac{2}{9}k)^{3/7} - 1$. $\square$

Lemma 2. *Suppose $h \geqslant 2$ and that for every finite set of positive real numbers B contined in some interval $[x, 2x]$, we have $|hB| + |B^h| \geqslant c|B|^{1+1/u}$. Then for any finite set A of positive real numbers, we have*

$$|E_h(A)| \geqslant \tfrac{c}{2}(ch^h h!/2)^{-\frac{1}{hu+1}} |A|^{1+\frac{h-1}{hu+1}}.$$

Proof. Let $k = |A|$ and break A into blocks

$$A_j = A \cap [2^{j-1}, 2^j) \qquad (j \in \mathbb{Z}).$$

Let

$$J = \{j : |A_j| > 0\},$$
$$m = \sum_{j \in J} |A_j|^{1+1/u}.$$

For each h-tuple of numbers $a_1, a_2, \ldots a_h \in A_j$, we have $\sum a_i \in [h2^{j-1}, h2^j)$ and $\prod a_i \in [2^{h(j-1)}, 2^{hj})$. Therefore, the sets hA_j are disjoint, as are the sets A_j^h. Hölder's inequality gives

$$k = \sum_{j \in J} |A_j| \leqslant |J|^{\frac{1}{u+1}} \left(\sum_{j \in J} |A_j|^{1+1/u} \right)^{\frac{u}{u+1}} = m^{\frac{u}{u+1}} |J|^{\frac{1}{u+1}},$$

which implies $|J| \geqslant k^{u+1} m^{-u}$. Choose one number a_j from each nonempty set A_j and set $n = 2 + [\frac{\log(h-1)}{\log 2}]$. For $0 \leqslant r \leqslant n-1$, let J_r be the subset of J with $j \equiv r \pmod{n}$. For some r, $|J_r| \geqslant \frac{|J|}{n}$. Form the set $C = \{a_j : j \equiv r \pmod{n}\}$. Since $a_{i+n} \geqslant 2^{n-1} a_i \geqslant h a_i$ for each i, the sums of distinct h-tuples of numbers in C are distinct. It follows from (2) and the hypothesis that

$$\begin{aligned} |E_h(A)| &\geqslant \max\left(\frac{1}{2} \sum_{j \in J} |hA_j| + |A_j^h|, \frac{|C|^h}{h!} \right) \\ &\geqslant \max\left(\frac{cm}{2}, \frac{k^{hu+h} m^{-uh}}{h^h h!} \right). \end{aligned}$$

The right side is minimized when $m^{hu+1} = 2k^{hu+h}/(ch^h h!)$, and this completes the proof. □

Combining Lemma 1 with Lemma 2 (taking $h = 2$, $c = \frac{7}{20}$, $u = 7$) gives Theorem 1. Theorem 2 follows from (4) and Lemmas 1 and 2. Proving $f_h(k) \gg k^{\beta(h)}$ with $\beta(h)$ tending to ∞ with h will require a non-trivial extension of Lemma 1 to the case $h \geqslant 3$, and it is not clear how this can be accomplished.

It is curious that nowhere in the argument was it necessary to assume the set A was a set of integers. Based on this observation, we make the following

Conjecture. *If A is a finite set of positive real numbers, then*

$$|E_h(A)| \gg_\varepsilon |A|^{h-\varepsilon}.$$

Before proving Theorem 3, we need a few definitions. A natural number n is said to be y-smooth if n is divisible by no prime factor $> y$. Denote by $\Psi(x,y)$ the number of y-smooth numbers $\leqslant x$. Important in the study of $\Psi(x,y)$ is the Dickman function $\rho(u)$, defined for $u \geqslant 0$ by

$$\begin{aligned} \rho(u) &= 1 \qquad (0 \leqslant u \leqslant 1), \\ \rho(u) &= 1 - \int_1^u \frac{\rho(v-1)}{v}\,dv \qquad (u > 1). \end{aligned}$$

We quote the following well-known results (Theorem 1.2 and Corollary 2.3 of [HiT]). Here we take $u = \frac{\log x}{\log y}$.

Lemma 3. *For any fixed $\varepsilon > 0$ we have*

$$\Psi(x,y) = x\rho(u)^{1+O(E(u))}$$

uniformly in the range

$$y \geqslant 2, 1 \leqslant u \leqslant y^{1-\varepsilon},$$

where

$$E(u) = \exp\{-(\log u)^{3/5-\varepsilon}\}.$$

Lemma 4. *Uniformly in $u \geqslant 3$, we have*

$$\rho(u) = \exp\left\{-u\left(\log u + \log_2 u - 1 + O\left(\frac{\log_2 u}{\log u}\right)\right)\right\}.$$

From now on assume h is fixed. In particular, constants implied by the $O-$ symbol may depend on h. Suppose x is large and set

$$\delta = \frac{2h\log h}{\log_2 x}, \qquad \alpha = \frac{h+\delta}{h-1}.$$

Let A be the set of $(\log x)^\alpha$-smooth numbers $\leqslant x$. Set $k = |A| = \Psi(x, (\log x)^\alpha)$ and $u = \frac{\log x}{\alpha \log_2 x}$. By Lemmas 3 (with $\varepsilon = \min(1/2, 1-1/\alpha)$) and 4, we have

$$\begin{aligned} k &= x\rho(u)^{1+O(E(u))} \\ &= x\exp\left\{-\frac{\log x}{\alpha\log_2 x}(\log_2 x - \log\alpha - 1) + O(L(x))\right\} \\ &= x^{1-1/\alpha}\exp\left\{O\left(\frac{\log x}{\log_2 x}\right)\right\}, \end{aligned}$$

where

$$L(x) = \frac{\log x \log_3 x}{(\log_2 x)^2}.$$

Consequently,

$$u = \frac{\log k}{(\alpha-1)\log_2 k}\left(1 + O\left(\frac{1}{\log_2 k}\right)\right). \tag{14}$$

Thus

$$|hA| \leqslant hx \leqslant k\rho(u)^{-1-O(E(u))}. \tag{15}$$

Lemma 3 also gives

$$|A^h| \leqslant \Psi(x^h, (\log x)^\alpha) = x^h \rho(hu)^{1+O(E(hu))} = k^h \left(\frac{\rho(hu)}{\rho(u)^h}\right)^{1+O(E(u))}. \tag{16}$$

By Lemma 4 and (14), we deduce

$$\begin{aligned}\rho(u) &= \exp\left\{-\frac{\log k}{\alpha-1} + \frac{1+\log(\alpha-1)}{\alpha-1}\frac{\log k}{\log_2 k} + O(L(k))\right\}\\ &\geqslant \exp\left\{-(h-1)(1-\delta+O(\delta^2))\log k - (h-1)\log h\frac{\log k}{\log_2 k} + O(L(k))\right\}\\ &\geqslant k^{-(h-1)}\exp\left\{h(h-1)\log h\frac{\log k}{\log_2 k} + O(L(k))\right\}.\end{aligned}$$

Similarly, we obtain

$$\frac{\rho(hu)}{\rho(u)^h} = \exp\left\{-h(h-1)\log h\frac{\log k}{\log_2 k} + O(L(k))\right\}.$$

Combining these estimates with (15) and (16) gives

$$|hA| + |A^h| \leqslant k^h \exp\left\{-h(h-1)\log h\frac{\log k}{\log_2 k} + O(L(k))\right\},$$

which completes the proof of Theorem 3.

References

[E] P. Erdös, *Problems and results on combinatorial number theory III*, Number Theory Day, New York 1976, Lecture Notes in Mathematics, vol. 626, Springer-Verlag, Berlin, 1977, pp. 43–72.

[ES] P. Erdös and Szemerédi, *On sums and products of integers*, Studies in Pure Mathematics, To the Memory of Paul Turán (P. Erdös, L. Alpár, G. Halász and A. Sárközy, eds.), Birkhäuser Verlag, Basel, 1983, pp. 213–218.

[F] G. Freiman, *Foundations of a structural theory of set addition*, Translations of Mathematical Monographs, vol. 37, Amer. Math. Soc., Providence, R. I., 1973.

[HaT] R. R. Hall and G. Tenenbaum, *Divisors*, Cambridge University Press, 1988.

[HiT] A. Hildebrand and G. Tenenbaum, *Integers without large prime factors*, J. Théor. Nombres Bordeaux **5** (1993), 411–484.

[JN] X. Jia and M. Nathanson, *Finite graphs and the number of sums and products*, preprint.

[N] M. Nathanson, *On sums and products of integers*, preprint.

[NT] M. Nathanson and G. Tenenbaum, *Inverse theorems and the number of sums and products*, preprint.

Department of Mathematics, University of Texas at Austin, Austin, Texas 78712

E-mail address: ford@math.utexas.edu