A QUESTION OF ERDŐS ON RECIPROCALS OF GAPS BETWEEN DIVISORS

DANIEL LARSEN

Abstract. Erdős conjectured that

$$\sum_{\substack{d_1,d_2|n\\ d_1<d_2}} \frac{1}{d_2-d_1} \ll 1 + \sum_{d_1,d_2 \text{ consecutive divisors of } n} \frac{1}{d_2-d_1}$$

remained bounded in the limit $n \to \infty$. Using a multiscale version of a construction of Tao [3] (taking n to be a product of nearly consecutive primes on different scales), we disprove this conjecture.

1. Introduction

For any positive integer n, we write $\operatorname{div}(n)$ for the set of divisors of n. For any set A of positive integers, we write

$$S(A) = \sum_{a_1,a_2 \text{ consecutive elements of } A} \frac{1}{a_2-a_1}$$

and

$$T(A) = \sum_{\substack{a_1,a_2\in A\\ a_1<a_2}} \frac{1}{a_2-a_1}.$$

We expect the ratio between $T(A)$ and $S(A)$ to measure some notion of regularity of A. If the gaps between consecutive elements of A are disordered, the consecutive gaps account for much of the "gap energy" under a harmonic weighting. Erdős [2] predicted that the divisors of a large integer should be disordered in this sense. In this note we show that this is not always the case.

Theorem 1.1 (Erdős #884 [1])**.** *The ratio*

$$\frac{T(\operatorname{div}(n))}{1+S(\operatorname{div}(n))} \tag{1}$$

is unbounded, where n ranges over the natural numbers.

Tao observed [3] that if the prime factors of n form a set with high additive energy (in some rough sense), this ratio could be large. In particular, in the impossible case that the prime factors form an arithmetic progression $\{a+kb\}_{k=1,\dots,l}$ whose length l is large compared to the size of the difference b,

$$T(\operatorname{div}(n)) \approx \sum_{1\le i<j\le l} \frac{1}{b(j-i)} \approx \frac{l\log l}{b}$$

whereas

$$S(\operatorname{div}(n)) \approx \sum_{1\le i<i+1=j\le l} \frac{1}{b(j-i)} = \frac{l-1}{b}.$$

(We expect that almost all the contribution in both cases come from gaps between prime divisors.) It makes little difference if the prime factors form a slightly perturbed arithmetic progression. In this case, there is no modular obstruction to l being large compared to b. However, in order to arrange this, Tao had to appeal to the k-tuples conjecture.

Our approach is to settle for a smaller value of l relative to b. As Tao pointed out, it then becomes possible to use an unconditional pigeonhole-based method to produce the primes. The only issue is that while the ratio between T and S is large, T is no longer large compared to 1. To overcome this difficulty, we repeat this construction on many difference scales, accumulating T-mass, while preserving the T to S ratio. There is a small issue of tracking cross-scale contributions to S, but some elementary analysis produces strong upper bounds on these contributions.

We use standard asymptotic notation and denote by $\omega(n)$ the number of prime divisors of n.

We would like to thank Terence Tao for his helpful comments on an earlier draft.

2. Construction

We begin with a slightly modified version of the key result from [3].

Proposition 2.1. *Suppose the prime factors of the square-free integer n have the form $t+h_1,\dots,t+h_K$, where the h_i form an $O(\log t)$-separated sequence of length $K \gg \log\log t$. Suppose further that $0 \le h_1 < \cdots < h_K < t$ and $(2(h_K-h_1))^K = o(t)$. Then*

$$\sum_{\substack{d_1,d_2 \text{ consecutive divisors of } n \\ \omega(d_1)=\omega(d_2)>1}} \frac{1}{d_2-d_1} = O\Big(\frac{1}{\log\log t}\Big) \tag{2}$$

Proof. Let $H = h_K - h_1$. Let $d_1 = \prod_{m=1}^{k}(t+h_{j_m})$, where the j-terms form an increasing sequence and write d_2 as a polynomial in t in the same way. If the polynomials differ by a non-constant term, then because the coefficients of both polynomials are bounded by $(2H)^K$, $d_2-d_1 \gg t$. The contribution to (2) from such terms can be bounded by $2^K/t = o(H^{-K})$. It remains to consider the contribution from pairs (d_1,d_2) whose corresponding polynomials differ only in their constant terms. By [3, Lemma 2.2],

$$d_2 - d_1 \gg \log t \prod_{m=1}^{k-1} (h_{j_k} - h_{j_m}). \tag{3}$$

There exist unique values $i_1,\dots,i_{k-1}$ such that $h_{j_k}-h_{j_m} \in [2^{i_m}, 2^{i_m+1})$. For a particular choice of k and $i_1,\dots,i_{k-1}$, there are $O(K)$ ways of picking j_k and $O(2^{i_m}/\log t)$ ways of picking every other j_m, taking advantage of the $\log t$-separation of the h_j. By (3), this choice of variables gives a contribution to (2) of

$$O\left(\frac{K\left(\prod_{m=1}^{k-1}(2^{i_m}/\log t)\right)}{(\log t)\prod_{m=1}^{k-1}2^{i_m}}\right) = o\big((1/\log t)^{k-1}\big).$$

Considering the $O(\log H)$ possible choices for each i_m,

$$\sum_{\substack{d_1,d_2 \text{ consecutive divisors of } n\\ \omega(d_1)=\omega(d_2)>1}} \frac{1}{d_2-d_1} \ll H^{-K} + \sum_{k=2}^{K} O\big((\log H/\log t)^{k-1}\big) = O\Big(\frac{\log H}{\log t}\Big).$$

The required bound follows because $\log H \ll \frac{\log t}{K}$. □

Corollary 2.2. *Under the same hypotheses as Proposition 2.1,*

$$\sum_{\substack{d_1,d_2 \text{ consecutive divisors of } n\\ \{d_1,d_2\}\not\subset\mathbb{P}}} \frac{1}{d_2-d_1} = O\Big(\frac{1}{\log\log t}\Big)$$

Proof. The additional summands compared to the sum in (2) satisfy $\omega(d_1)\neq\omega(d_2)$. The bounds $t^{\omega(d_1)} < d_1 < (2t)^{\omega(d_1)}$ and $t^{\omega(d_2)} < d_2 < (2t)^{\omega(d_2)}$, along with the fact $2^K = o(t)$, imply that $\omega(d_2) > \omega(d_1)$ and, further, that $d_1 = o(d_2)$. The additional contribution to the sum is therefore bounded, up to a constant factor, by

$$\sum_{1<d|n} \frac{1}{d} < \exp\Big(\sum_{p|n}\frac{1}{p}\Big) - 1 \ll \frac{K}{t}.$$

□

Let $f(x) = \log^* x$ be the number of times the natural logarithm function must be iterated to obtain a value ≤ 1.

Lemma 2.3. *For any large integer t there exists a set of primes $p_1,\dots,p_l$ between t and $2t$ with the property that $(p_i - t)_{i=1,\dots,l}$ is an increasing $O(\log t)$-separated sequence of length $\left\lfloor \frac{\log t}{f(t)\log\log t}\right\rfloor$ within an interval of length $O(\frac{\log^2 t}{f(t)\log\log t})$, where the implied constants are absolute.*

Proof. By the prime number theorem, there are $> t/(2\log t)$ primes between t and $2t$. Let

$$\mathfrak{S}(n) = \prod_{p|n}\Big(1-\tfrac{1}{p}\Big)^{-1}\prod_{p\nmid n}\left(1-\frac{2}{p}\right)\left(1-\frac{1}{p}\right)^{-2}.$$

Then,

$$\sum_{n\leq N}\mathfrak{S}(n) \ll \sum_{n\leq N}\prod_{p|n}\left(1+\frac{1}{p}\right) \leq \sum_{n\leq N}\sum_{d|n}\frac{1}{d} = \sum_{d\leq N}\sum_{n\in d\mathbb{N}\cap[1,N]}\frac{1}{d} \leq \sum_{d\leq N}\frac{N}{d^2} \ll N.$$

Combining this bound with the Brun sieve, there exists an absolute constant C such that for any integer N,

$$\#\{(p_1,p_2) : p_1,p_2\in[t,2t],\ 0<p_2-p_1<N\} < \frac{CNt}{\log^2 t}. \tag{4}$$

Let ε be sufficiently small in terms of C. Set $N = \varepsilon\log t$,

$$K = \Big\lfloor \frac{\log t}{f(t)\log\log t}\Big\rfloor,$$

and $H = 8K\log t$.

Divide $[t, 2t]$ into consecutive intervals I_j of length H each, discarding the last (partial) interval. Let $J_{\text{good}} = \{j : |I_j \cap \mathbb{P}| \geq 2K\}$ and $J_{\text{bad}} = \{j : |I_j \cap \mathbb{P}| < 2K\}$. Let $\mathcal{P}_{\text{good}} = \bigcup_{j \in J_{\text{good}}}(I_j \cap \mathbb{P})$ and $\mathcal{P}_{\text{bad}} = \bigcup_{j \in J_{\text{bad}}}(I_j \cap \mathbb{P})$. Since

$$|\mathcal{P}_{\text{bad}}| \leq \frac{t}{H} \cdot 2K = \frac{t}{4 \log t},$$

we have $|\mathcal{P}_{\text{good}}| > t/(4 \log t)$. Letting k_j denote the number of prime gaps within I_j less than N,

$$\sum_{j \in J_{\text{good}}} k_j \;\leq\; \frac{CNt}{\log^2 t} \;<\; \frac{t}{10 \log t} \;<\; \frac{|\mathcal{P}_{\text{good}}|}{2} - |J_{\text{good}}|,$$

so most intra-interval prime gaps in $\mathcal{P}_{\text{good}}$ are $\geq N$.

The pigeonhole principle then implies there exists $j \in J_{\text{good}}$ such that most prime gaps in $I_j \cap \mathbb{P}$ are $\geq N$, giving an N-separated subset of $I_j \cap \mathbb{P}$ of size $\geq |I_j \cap \mathbb{P}|/2 \geq K$. □

Corollary 2.4. *For any large integer t, there exists a set B of primes between t and $2t$ such that defining n to be the product of the elements of B, we have $S(\operatorname{div}(n)) = o(T(\operatorname{div}(n)))$. Moreover, B can be chosen so that*

$$T(B) \gg \frac{1}{f(t)}, \tag{5}$$

$|B| = o(\frac{\log t}{\log \log t})$, and the diameter of B is less than $\log^2 t$.

Proof. Take B to be a subset of size K as produced in the proof of Lemma 2.3. As B is a set of integers of size K which has no two elements within distance N of each other and whose diameter is at most H, we have

$$S(B) \leq \frac{K-1}{N} \ll \frac{K^2}{H}$$

and by [3, (1.2)],

$$T(B) \gg \frac{K^2 \log K}{H} \gg \frac{1}{f(t)}.$$

By Corollary 2.2, the remaining contribution to $S(\operatorname{div}(n))$ is $O(1/\log \log t)$. Thus,

$$\frac{T(\operatorname{div}(n))}{S(\operatorname{div}(n))} \geq \frac{T(B)}{S(B) + O(1/\log \log t)} \gg \frac{\log \log t}{f(t)},$$

which diverges. □

Let t_0 be large, and let $t_i = \exp(t_{i-1})$ for $i \geq 1$.

Proposition 2.5. *Suppose for each t_i, there exists a set of primes B_i between t_i and $2t_i$ such that*

$$S(\operatorname{div}(m_i)) = o(T(\operatorname{div}(m_i))), \tag{6}$$

where m_i is the product of the elements of B_i. Suppose further that $k_i = o(\frac{\log t_i}{\log \log t_i})$ and $k_i w_i \ll \sqrt{t_i}$, where k_i is the size of B_i and w_i is the diameter of B_i. Then $S(\operatorname{div}(n_i)) = o(T(\operatorname{div}(n_i)))$, where $n_i = \prod_{j \leq i} m_j$.

Proof. Let $n = n_r$ for some large r. For a divisor $d \mid n$, we write $d = \prod_i d^i$, where d^i denotes the product of all prime factors of d which lie in B_i. Let $d_1, d_2 \in \operatorname{div}(n)$ such that for some i, $\omega(d_1^i) \neq \omega(d_2^i)$. Take i to be the largest such value. Let

$$\Delta_j = \log d_2^j - \log d_1^j.$$

Without loss of generality we can assume $\Delta_i > 0$. On one hand, by assumption, for all $j > i$, d_1^j and d_2^j have the same number of prime factors, so that

$$|\Delta_j| \le k_j \log(1 + w_j/t_j) \ll t_j^{-1/2}. \tag{7}$$

On the other hand, for $j < i$, $|\Delta_j| \le k_j \log(2t_j)$ by the trivial bound. Meanwhile,

$$\Delta_i \ge \log t_i - k_i \log(1 + w_i/t_i).$$

Therefore

$$\log d_2 - \log d_1 \gg \log t_i. \tag{8}$$

Let X be the set of pairs (d_1, d_2), with $d_1 < d_2$ consecutive elements of $\operatorname{div}(n)$. Let Y be the subset of $(d_1, d_2) \in X$ such that

$$\omega(d_1^i) = \omega(d_2^i)$$

for all i. Let $Z = X \setminus Y$. Stratify Y into $Y_{i,\ell}$, the subset of pairs $(d_1, d_2) \in Y$ such that i is the smallest value for which $d_1^i \neq d_2^i$ and ℓ is the largest value for which $d_1^\ell d_2^\ell > 1$. Stratify Z into $Z_{i,\ell}$ where i is the largest value such that $\omega(d_1^i) \neq \omega(d_2^i)$ and ℓ is the largest value for which $d_1^\ell d_2^\ell > 1$.

By (8),

$$\sum_{i \le \ell} \sum_{(d_1,d_2) \in Z_{i,\ell}} \frac{1}{d_2 - d_1} \ll \sum_{i \le \ell} \frac{|Z_{i,\ell}|}{t_\ell} \le \frac{2^{2\omega(n_\ell)}}{t_\ell} = t_\ell^{o(1)-1}. \tag{9}$$

Now suppose $(d_1, d_2) \in Y_{i,\ell}$, with $\ell > i$. Then

$$|\Delta_i| \ge \log\left(1 + \frac{1}{(2t_i)^{k_i}}\right),$$

so considering (7), $|\log d_1 - \log d_2| \gg (2t_i)^{-k_i}$. Hence,

$$\sum_{i < \ell} \sum_{(d_1,d_2) \in Y_{i,\ell}} \frac{1}{d_2 - d_1} \ll \sum_{i < \ell} \frac{|Y_{i,\ell}| \cdot (2t_i)^{k_i}}{t_\ell} = t_\ell^{o(1)-1}, \tag{10}$$

using the fact that $t_i \le \log t_\ell$ and the hypothesis that $k_i = o(\frac{\log t_i}{\log\log t_i})$. Finally,

$$\begin{aligned}\sum_{(d_1,d_2) \in Y_{i,i}} \frac{1}{d_2 - d_1} &\le \left(\sum_{d \mid n_{i-1}} \frac{1}{d}\right) \sum_{d_1, d_2 \text{ consecutive divisors of } m_i} \frac{1}{d_2 - d_1} \\ &< 2S(\operatorname{div}(m_i)).\end{aligned} \tag{11}$$

Combining (9), (10), and (11), we obtain

$$S(\operatorname{div}(n)) = \sum_{(d_1,d_2) \in X} \frac{1}{d_2 - d_1} \ll \sum_{\ell \le r} \left(t_\ell^{o(1)-1} + S(\operatorname{div}(m_\ell))\right) \ll \sum_{\ell \le r} S(\operatorname{div}(m_\ell)),$$

whereas

$$T(\operatorname{div}(n)) > \sum_{\ell \le r} T(\operatorname{div}(m_\ell)).$$

The proposition follows by applying (6). ∎

Let B_i be as in Corollary 2.4 for $t = t_i$. By (5),

$$T(\mathrm{div}(n_i)) \gg \sum_{j \le i} \frac{1}{f(t_j)} = \sum_{j \le i} \frac{1}{j + \log^* t_0} \to \infty, \tag{12}$$

as $i \to \infty$. Combining this with Proposition 2.5, Theorem 1.1 follows.

References

[1] T. F. Bloom, `https://www.erdosproblems.com`

[2] P. Erdős, *Some of my new and almost new problems and results in combinatorial number theory*, Number theory: diophantine, computational, and algebraic aspects: proceedings of the international conference held in Eger, Hungary, July 29–August 2, 1996, 1998, p. 169–180.

[3] T. Tao, *On the sum of reciprocals of gaps between divisors*, preprint.