LARGE SUBSUMS OF THE MÖBIUS FUNCTION

IMRE Z. RUZSA

Abstract. We find large subsums of the Möbius function for divisors of an integer.

To the memory of Eduard Wirsing.

1. Introduction

Put

$$M(n,t) = \sum_{d|n, d\le t} \mu(d),$$

where μ is the Möbius function. We shall consider

$$f(n) = \max_t |M(n,t)|$$

and

$$g(n) = M(n, \sqrt{n})$$

for suqarefree values of n. The reason of this restriction for f is that clearly $M(n,t) = M(n',t)$, where n' is the product of prime divisors of n. For squarefree n we have

$$M(n,t) = -\mu(n) M(n, n/t),$$

(unless $t|n$ when there is an extra $\mu(t)$), that is, this funcion exhibits a symmetry around $\sqrt{n}$, in particular, $g(n) = 0$ if $\mu(n) = 1$. This makes this turning point relevant and the estimation of g somewhat more difficult. For not squarefree n there is nothing special in the stopping point $\sqrt{n}$.

Clearly

$$|g(n)| \le f(n) \le 2^{\omega(n)}, \tag{1.1}$$

where ω denotes the number of prime divisors. Since

$$\max_{n\le x} \omega(n) \sim (\log x)/\log\log x,$$

we conclude that

$$\max_{n\le x} |g(n)| \le \max_{n\le x} f(n) \le 2^{(1+o(1))(\log x)/\log\log x}. \tag{1.2}$$

We consider how sharp are these inequalities.

Theorem 1.

$$\max_{n\le x} f(n) > 2^{(1-o(1))(\log x)/\log\log x}.$$

Theorem 2.

$$\max_{n\le x} |g(n)| > 2^{(2/3-o(1))(\log x)/\log\log x}.$$

1991 *Mathematics Subject Classification.* 11P32, 11N99.
Supported by NKFI grants K-129335, K-119528, KKP-133819.

Conjecture.

$$\max_{n\le x} |g(n)| > 2^{(1-o(1))(\log x)/\log\log x}.$$

We give some heuristics to support this conjecture in the next section.

It is worthwhile to recall that the typical values of these functions are very small. With the notation

$$\delta(t) = \mathrm{d}\{n : M(n,t) \ne 0\}$$

Erdős and Hall[2] proved that

$$\delta(t) = O\left((\log t)^{-c}\right),\ c = 0.0579\ldots.$$

Their proof can easily be adapted to show that

$$\mathrm{d}\left\{n : g(n) = 0\right\} = 1.$$

Probably the estimate

$$|\{n \le x : g(n) \ne 0\}| < x/(\log x)^c$$

also holds with some positive c, but I have not checked the details

Maier[4] proved that typically $f(n)$ is between two powers of $\log\log n$. The best exoponents, due to Maier-Tenenbaum[5] are that for almost all n

$$(\log\log n)^{c_1-\varepsilon} < f(n) < (\log\log n)^{c_2+\varepsilon},$$

$c_1 = 0.28754...$, $c_2 = \log 2$.

For the square mean, de la Bretèche, Dress, Tenenbaum[6] proved that

$$\frac{1}{x}\sum_{n\le x} |M(n,t)|^2 = L + o(1)$$

with some constant L *uniformly* as long as $t \to \infty$ and $t = o(x)$.

Conjecture.

$$\sum_{n\le x} g(n)^2 = Lx + o(x).$$

I can prove the weaker estimate

$$\sum_{n\le x} g(n)^2 = O(x).$$

2. Proof of Theorem 1

Proof. We use the Dirichlet series of the Möbius function for divisors of n at imaginary values:

$$D(\alpha) = \sum_{d|n} \mu(d) d^{i\alpha} = -i\alpha \int_1^n M(n,t) t^{i\alpha-1}\, dt,$$

whence

$$|D(\alpha)| < |\alpha| \int_1^n \frac{|M(n,t)|}{t}\, dt < |\alpha| f(n) \log n.$$

So to find a large value of $f(n)$ it is sufficient to find a single large value of D for a not too large α.

Write $D(\alpha)$ as an Euler product:

$$D(\alpha) = \prod_{p|n} \left(1 - p^{i\alpha}\right) = n^{i\alpha/2} \prod_{p|n} \left(p^{-i\alpha/2} - p^{i\alpha/2}\right) = n^{i\alpha/2}(-2)^k \prod_{p|n} \sin \frac{\alpha \log p}{2}$$

with $k = \omega(n)$, so

$$|D(\alpha)| = 2^k \prod_{p|n} \left|\sin \frac{\alpha \log p}{2}\right|.$$

For n we shall take the product of primes of an interval $(u, 2u)$ and put $\alpha = \pi/(\log u)$. For each prime we have

$$\frac{\pi}{2} < \frac{\alpha \log p}{2} < \frac{\pi}{2}\left(1 + \frac{\log 2}{\log u}\right) = \frac{\pi}{2} + \varepsilon$$

with $\varepsilon = (\log 2)/\log u$, hence each sinus is $> 1 - \varepsilon$ and $|D(\alpha)| > 2^{(1-\varepsilon)k}$.

The product of primes in the interval $(u, 2u)$ is

$$\prod_{u<p<2u} p \le e^{\Psi(2u)-\Psi(u)} = e^{u(1+o(1))},$$

so we can take $u \sim \log x$. The number of primes in this interval is $k \sim (\log x)/\log\log x$. □

Let n be an integer with large $f(n)$ and $\mu(n) = 1$, and p a prime, $p \nmid n$. We have

$$M(pn, t) = M(n, t) - M(n, t/p) = M(n, t) + M(pn/t),$$

(unless $p|t$ and $(t/p)|n$), hence

$$g(pn) = 2M(n, \sqrt{pn}).$$

So if we could find a prime such that $\sqrt{pn}$ is near to the value maximizing $M(n, t)$, we would have a large value of g. Unfortunately the above argument does not help to locate the place of maximum.

It is possible to express $g(n)$ by $D(\alpha)$ through a Perron formula, but it is not obvious how to use this to estimate $g(n)$.

Problem. Let p_i be the i-th prime, and

$$p_1 p_2 \dots p_r \le x < p_1 p_2 \dots p_{r+1}.$$

Then $\max_{n\le x} \omega(n) = r$. From the theorem of Erdős[1] we know that $f(n) < c2^r/\sqrt{r}$. The proof above gives $f(n) > 2^{r - cr/\log r}$. Where lies the truth?

3. Prolegomena to Theorem 2

Consider $n = p_1 p_2 \dots p_k$, $k = 2m + 1$, with primes p_i such that $p_1 < \dots < p_k$ and $p_1 \dots p_{m+1} > p_{m+2} \dots p_{2m+1}$. Then exactly those divisors are $< \sqrt{n}$ which are composed of at most m primes. Consequently

$$g(n) = \sum_{j=0}^{m} (-1)^j \binom{k}{j} = (-1)^m \binom{k-1}{m} \sim \pm c2^k/\sqrt{k}.$$

In terms of k this is best possible. Erdős[1] proved that whenever $\omega(n) = k$, we have

$$f(n) \le \binom{k-1}{\lfloor (k-1)/2 \rfloor}.$$

(For even k we have $g(n) = 0$ but $f(n)$ can be that large.)

Try to use the primes in an interval $(u, u+v)$. This certainly works if $u^{m+1} > (u+v)^m$, that is,

$$\left(1 + \frac{v}{u}\right)^m < u.$$

As $\left(1 + \frac{v}{u}\right)^m \approx e^{mv/u}$, this is about the same as $mv < u \log u$.

We need $k = 2m + 1$ primes. The average distance between primes around u is $\log u$. This means $v \sim 2m \log u$, or $u > 2m^2 \sim k^2/2$. To ensure $n < x$ the condition is $(k^2/2)^k < x$, which yields

$$k \sim \frac{1}{2} \frac{\log x}{\log \log x}.$$

This simple construction gives Theorem 2 with $1/2$ in the place of $2/3$.

The improvement, to be detailed in the next section, is based on the observation that we do not actually need that *all* divisors composed of at most m primes should be $< \sqrt{n}$, only that most of them be.

4. Proof of Theorem 2

Write $n = p_1 \dots p_k$, $k = 2m + 1$, $u = p_1 < \dots < p_k = u + v$. Our choice will be

$$k \sim \left(\frac{2}{3} - \varepsilon\right) \frac{\log x}{\log \log x}$$

and $u \sim k^{3/2} \log k$. Around u the typical distance between primes is $\log u$, we expect $v \sim k \log u$. To show that this is indeed the case we recall Heath-Brown's estimate for primes in intervals [3] by which $\pi(v + w) - \pi(v) \sim w/\log w$ as long as $u^{7/12+\varepsilon} < w < u$ and in our case the interval is longer than $u^{2/3}$.

We estimate the ratio of products of m terms below $\sqrt{n}$. Put $a = (\log n)/k$, $\log p_i = a + b_i$. Clearly

$$\log u < a < \log(u + v).$$

For $d|n$, $\omega(d) = j$ we have

$$\log d = ja + \text{ sum of some } b_i.$$

If we can ensure that from the 2^k sum formed by the numbers b_i only $o(2^k/\sqrt{k})$ are $> a/2$, then

$$g(n) = (-1)^m \binom{k-1}{m} + o\left(\frac{2^k}{\sqrt{k}}\right) \sim \pm c 2^k/\sqrt{k}$$

will hold as wanted.

We expect the sum of the b_i to have approximately a normal distribution. To make this exact we use Bernstein's inequality in the following form. Let $\xi_1, \dots, \xi_k$ be independent random variables with zero mean and $|\xi_i| \le M$. Then

$$\text{Prob}\left(\sum \xi_i \ge t\right) \le \exp \frac{-t^2/2}{\sum \mathbb{E}(\xi_i^2) + Mt/3},$$

see e.g. https://en.wikipedia.org/wiki/Bernstein_inequalities_(probability_theory). We apply this for ξ_i assuming the values $\pm b_i/2$ with probability $1/2$ each. (Note that

$\sum b_i = 0$, hence $\sum_{i\in I} b_i = (1/2)\left(\sum_{i\in I} b_i - \sum_{i\notin I} b_i\right)$, so subsums of b_i have the same distribution as $\sum \xi_i$). Clearly

$$|b_i| < \log\frac{u+v}{u} < v/u,$$

so we can set $M = v/(2u)$. We put $t = a/2$, and estimate $\mathbb{E}(\xi_i^2)$ by M^2.

We claim that with this choice of the arguments (for sufficiently large value of the parameters) we have

$$\frac{t^2/2}{\sum \mathbb{E}(\xi_i^2) + Mt/3} > \log k,$$

hence the probability of large values will be $< 1/k$, that is, the number of subsums with an unusual large contribution from the b_i is $O(2^k/k)$.

After substituting $t = a/2$, and estimating $\mathbb{E}(\xi_i^2)$ by M^2 we obtain that it is sufficient to show

$$8\log k\left(kM^2 + aM/6\right) < a^2. \tag{4.1}$$

To see this we express these quantities asymptotically by k. We have

$$u \sim k^{3/2}\log k,\ v \sim k\log u \sim \frac{3}{2}k\log k,$$

$$M = \frac{v}{2u} \sim \frac{3}{4\sqrt{k}},$$

hence

$$8(\log k)kM^2 \sim \frac{9}{2}\log k = o(a^2),$$

as $a > \log u > \log k$, so the first summand in (4.1) is small enough. For the second observe that

$$M\log k \sim \frac{3\log k}{4\sqrt{k}} = o(1),$$

so $aM\log k = o(a^2)$ as well. This concludes the proof of (4.1).

References

[1] P. Erdős, *On a problem in elementary number theory*, Math. Student (1949).
[2] P. Erdős and R. R. Hall, *On the Möbius function*, J. Reine Angew. Math. **315** (1980), 121–128.
[3] D.R. Heath-Brown, *The number of primes in a short interval*, J. Reine Angew. Math. (1989).
[4] H. Maier, *On the Möbius function*, Trans. Amer. Math. Soc. **301** (1987), 649–664.
[5] H. Maier and G. Tenebaum, *On the normal concentration of divisors 2*, Math. Proc. Cambridge Phil. Soc. **147** (2009), 593–614.
[6] G. Tenenbaum R. de la Bretèche, F. Dress, *Remarques sur une somme liée à la fonction de Möbius*, Mathematika (2020).

Alfréd Rényi Institute of Mathematics, Budapest, Pf. 127, H-1364 Hungary
E-mail address: ruzsa@renyi.hu