SMOOTH NUMBERS IN SHORT INTERVALS

K. SOUNDARARAJAN

Abstract. Assume the Riemann Hypothesis. For every $\epsilon > 0$ we show that there is a constant $C(\epsilon)$ such that for all large x, the interval $[x, x + C(\epsilon)\sqrt{x}]$ contains an integer all of whose prime factors are less than x^ϵ.

A natural number n is called y-smooth if all its prime factors are below y. We let $\mathcal{S}(y)$ denote the set of y-smooth numbers, and let $\Psi(x, y)$ denote the number of such integers below x. If we write $y = x^{\frac{1}{u}}$ then it is known that $\Psi(x, y) \sim \rho(u)x$ where $\rho(u)$ denotes the Dickman function defined by $\rho(u) = 1$ for $0 \le u \le 1$ and for $u \ge 1$ is defined as the unique continuous solution to the differential-difference equation $u\rho'(u) = -\rho(u-1)$. This asymptotic formula was published first by Dickman [3] for fixed values of u and as $x \to \infty$; recently Soundararajan [14] has pointed out that such an asymptotic formula may be found in Ramanujan's unpublished papers. Later work has established asymptotic formulae for $\Psi(x, y)$ uniformly for u in a wide range; see for example the surveys [5] and [9].

In this note we are concerned with the existence of smooth numbers in short intervals. For a wide range of the variables x, y, and z, it is expected that

$$\Psi(x + z, y) - \Psi(x, y) \asymp \frac{z}{x}\Psi(x, y).$$

It is also of interest to establish the existence of smooth numbers in such short intervals, even if one is not able to exhibit a positive proportion of such numbers. One motivation for this problem is the analysis of Lenstra's elliptic curve factorization algorithm [11] (and see also [13]) where one wishes to find integers in $[x, x + 4\sqrt{x}]$ which are $\exp(\sqrt{\log x \log\log x})$ smooth.

Regarding this problem, an important advance was made by Balog [1] who showed that for any fixed $\epsilon > 0$ and x large, the interval $[x, x + x^{\frac{1}{2}+\epsilon}]$ contains many x^ϵ-smooth integers. Harman [7] has obtained a strengthening of this result, allowing ϵ to be a function of x. The problem for intervals of length of size $\sqrt{x}$ has proved resistant, and here Harman [8] has established that the interval $[x, x + \sqrt{x}]$ contains integers that are $x^{\frac{1}{4\sqrt{e}}}$-smooth. If one expands the interval a little to consider $[x, x + C\sqrt{x}]$ for some constant $C > 0$, then recently Matomäki [12], advancing an approach of Croot [2], has shown that such intervals (for a suitably large value of C) contain $x^{\frac{1}{5\sqrt{e}}+\epsilon}$-smooth numbers.

The author is supported in part by a grant from the National Science Foundation.

One may wonder if the Riemann Hypothesis is of use in this problem. Assuming RH, Xuan [15] has shown that intervals $[x, x + x^{\frac{1}{2}}(\log x)^{1+\epsilon}]$ contain x^ϵ-smooth integers. Recently Ganguly and Pal [4] have noted that if, in addition to RH, one assumes a strong conjectural estimate for $\pi S(t)$ (which is the argument of $\zeta(1/2+it)$), then intervals $[x, x+x^{\frac{1}{2}}(\log x)^{\frac{1}{2}+\epsilon}]$ contain x^ϵ-smooth numbers. We improve upon Xuan's work by establishing the following theorem, which unfortunately is still not strong enough to be applicable to the analysis of Lenstra's algorithm.

Theorem. *Assume the Riemann Hypothesis. Let x be large and suppose that $x \geq y \geq \exp(5\sqrt{\log x \log\log x})$, and write $y = x^{\frac{1}{u}}$. There is an absolute constant B such that with $z = Bu\sqrt{x}/\rho(u/2)$ we have*

$$\Psi(x+z, y) - \Psi(x, y) \gg_\epsilon zx^{-\epsilon}.$$

By using estimates for divisor functions in short intervals we can obtain a better lower bound for the number of smooth integers in short intervals, but our methods would not give a positive proportion. Our Theorem sheds no light on y-smooth integers with y smaller than $\exp(\sqrt{\log x \log\log x})$, and it would be interesting to devise alternative approaches in this regime. Our Theorem is also likely to be very far from the truth about smooth numbers in short intervals. For example, one would expect that for every $\epsilon > 0$ there exists a constant $C(\epsilon)$ such that every interval $[x, x + C(\epsilon)\log x]$ contains an x^ϵ-smooth number. This would be analogous to Cramér's conjecture on the distribution of prime numbers, and note that a similarly large gulf exists between what can be established about primes on RH and the expected truth. As with primes, one can say more about the existence of smooth numbers in almost all short intervals; this problem has been considered by Hafner [6] but his work remains unpublished.

We now turn to the proof of our Theorem. Let the parameters x, y, z and u be as in the Theorem, and define δ by $xe^{2\delta} = x+z$. Let

$$M(s) = \sum_{\substack{\sqrt{x}y^{-1/3} \leq n \leq \sqrt{x}y^{-1/4} \\ n \in \mathcal{S}(y)}} \frac{1}{n^s}.$$

Our proof of the Theorem is based upon considering

$$I = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} -\frac{\zeta'}{\zeta}(s)M(s)^2 x^s \frac{(e^{\delta s}-1)^2}{s^2} ds$$

where $c = 1 + \frac{1}{\log x}$.

By shifting contours to the left if $\xi > 1$ and to the right if $\xi \leq 1$ we may see that

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \xi^s \frac{ds}{s^2} = \begin{cases} \log \xi & \text{if } \xi \geq 1 \\ 0 & \text{if } 0 < \xi \leq 1. \end{cases}$$

Therefore

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \xi^s \frac{(e^{\delta s}-1)^2}{s^2} ds = \begin{cases} \min(\log(e^{2\delta}\xi), \log(1/\xi)) & \text{if } e^{-2\delta} \leq \xi \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Hence

$$I = \sum_{x\le n\le xe^{2\delta}} \sum_{\substack{n=rm_1m_2 \\ m_1,m_2\in\mathcal{S}(y) \\ \sqrt{x}y^{-1/3}\le m_1,m_2\le\sqrt{x}y^{-1/4}}} \Lambda(r)\min\Big(\log\frac{e^{2\delta}x}{n},\log\frac{n}{x}\Big).$$

Note that $r = n/(m_1m_2)$ is at most y, and so the integers n counted in the RHS above are all y-smooth. Moreover the number of ways of writing n as rm_1m_2 is at most $d_3(n) \ll x^{\epsilon}$. Therefore we conclude that

$$I \ll \delta x^{\epsilon}(\Psi(x+z,y)-\Psi(x,y)). \tag{1}$$

We shall now derive a lower bound for I which will prove the Theorem. We move the line of integration in the definition of I to the line $\mathrm{Re}(s) = -\frac{1}{2}$. We encounter poles at $s = 1$ and at the non-trivial zeros $\rho = \frac{1}{2}+i\gamma$ of $\zeta(s)$. Thus we find that I equals

(2)

$$x(e^{\delta}-1)^2M(1)^2-\sum_{\rho}M(\rho)^2x^{\rho}\Big(\frac{e^{\delta\rho}-1}{\rho}\Big)^2-\frac{1}{2\pi i}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty}\frac{\zeta'}{\zeta}(s)M(s)^2x^s\frac{(e^{\delta s}-1)^2}{s^2}ds.$$

Using the functional equation for $\zeta(s)$ and Stirling's formula, we obtain that $|\frac{\zeta'}{\zeta}(-\frac{1}{2}+it)| \ll \log(2+|t|)$. Since $(e^{\delta s}-1)^2/s^2 \ll \min(\delta^2, 1/|s|^2)$ for all complex s with $-\frac{1}{2}\le \mathrm{Re}\ (s)\le 2$, we find that the integral appearing in (2) is bounded by

$$\ll x^{-\frac{1}{2}}\int_{-\infty}^{\infty}|M(-\tfrac{1}{2}+it)|^2\log(2+|t|)\min\Big(\delta^2,\frac{1}{1/4+t^2}\Big)dt.$$

We now split the interval $(-\infty,\infty)$ into the sets $\mathcal{I}_0 = \{|t|\le 1/\delta\}$, and $\mathcal{I}_j = \{2^{j-1}/\delta \le |t| \le 2^j/\delta\}$ for $j\in\mathbb{N}$. By appealing to a standard mean-value theorem for Dirichlet polynomials (see for example Theorem 9.1 of [10]) we find that the contribution from $t\in\mathcal{I}_0$ is

$$\begin{aligned}
&\ll x^{-\frac{1}{2}}\delta^2\log(1/\delta)\int_{\mathcal{I}_0}|M(-\tfrac{1}{2}+it)|^2dt\\
&\ll x^{-\frac{1}{2}}\delta^2\log x\Big(\frac{\sqrt{x}}{y^{\frac{1}{4}}}+\frac{1}{\delta}\Big)\sum_{\substack{n\in\mathcal{S}(y)\\ \sqrt{x}/y^{\frac{1}{3}}\le n\le\sqrt{x}/y^{\frac{1}{4}}}} n\\
&\ll \delta^2\log x\frac{\sqrt{x}}{y^{\frac{1}{2}}}\Big(\frac{\sqrt{x}}{y^{\frac{1}{4}}}+\frac{1}{\delta}\Big)M(1).
\end{aligned}$$

Similarly we find that for $j\ge 1$ the contribution from the interval $\mathcal{I}_j$ is

$$\ll \frac{j\delta^2\log x)}{2^j}\frac{\sqrt{x}}{\sqrt{y}}\Big(\frac{\sqrt{x}}{y^{\frac{1}{4}}}+\frac{1}{\delta}\Big)M(1).$$

We conclude that the integral in (2) is bounded by

$$\ll \delta^2\log x\frac{\sqrt{x}}{\sqrt{y}}\Big(\frac{\sqrt{x}}{y^{\frac{1}{4}}}+\frac{1}{\delta}\Big)M(1). \tag{3}$$

Now we turn to the sum over zeros in (2). This sum is bounded by

$$\ll x^{\frac{1}{2}} \sum_{\gamma} |M(\tfrac{1}{2}+i\gamma)|^2 \min\Big(\delta^2, \frac{1}{1/4+\gamma^2}\Big).$$

To estimate this, we decompose the sum into cases depending on whether $\gamma \in \mathcal{I}_j$ with $\mathcal{I}_j$ as earlier. We shall prove that the contribution from the zeros in $\mathcal{I}_j$ for any $j \geq 0$ is

$$\ll \sqrt{x}\frac{j+1}{2^j}\delta^2 \log(1/\delta)\Big(\frac{\sqrt{x}}{y^{\frac{1}{4}}} + \frac{1}{\delta}\Big)M(1). \tag{4}$$

Summing over all j, it then follows that the sum over zeros in (2) is

$$\ll \sqrt{x}\delta^2 \log(1/\delta)\Big(\frac{\sqrt{x}}{y^{\frac{1}{4}}} + \frac{1}{\delta}\Big)M(1). \tag{5}$$

Note that the contribution in (3) is dominated by that in (5). Thus, combining (2), (3) and (5) we conclude that

$$I \geq x\delta^2 M(1)^2 - A\sqrt{x}\delta^2 \log x\Big(\frac{\sqrt{x}}{y^{\frac{1}{4}}} + \frac{1}{\delta}\Big)M(1),$$

for an appropriate absolute constant A. Since $\rho(u/2) = u^{-u/2(1+o(1))}$, in our range of δ and y we have that $1/\delta \geq \sqrt{x}/y^{\frac{1}{4}}$. Moreover, using the asymptotic formula for smooth numbers, we see that $M(1) \geq \rho(u/2)(\log y)/24$. Choosing B suitably in terms of A, from the above remarks we find that $I \geq x\delta^2 M(1)^2/2$, and by (1) the Theorem follows.

It remains lastly to justify the bound (4). We treat the case $j = 0$, the other cases being similar. The proof is entirely standard and we sketch the details quickly; indeed we could obtain asymptotic formulae for such sums but we do not need this. Let $\xi(s) = s(s-1)\pi^{-s/2}\Gamma(s/2)\zeta(s)$ denote Riemann's ξ-function which is entire, and whose zeros are the non-trivial zeros of the $\zeta(s)$. We consider, with $c = 1 + 1/\log x$,

$$J := \frac{1}{2\pi i}\int_{(c)} \frac{\xi'}{\xi}(s)M(s)M(1-s)\Big(\frac{e^{\delta(s-1/2)} - e^{-\delta(s-1/2)}}{(s-1/2)}\Big)^2 ds.$$

We now move the line of integration to $\mathrm{Re}(s) = 1-c$. There are poles at the non-trivial zeros of $\zeta(s)$ and these contribute

$$\sum_{\gamma} |M(1/2+i\gamma)|^2 \Big(\frac{2\sin(\delta\gamma)}{\gamma}\Big)^2.$$

To handle the integral on the line $\mathrm{Re}(s) = 1-c$ we use the functional equation $\xi'/\xi(s) = -\xi'/\xi(1-s)$ and then make a change of variable $w = 1-s$. In this manner we recognize the integral on $\mathrm{Re}(s) = 1-c$ as being $-J$. Thus we find that

$$2J = \sum_{\gamma} |M(1/2+i\gamma)|^2 \Big(\frac{2\sin(\delta\gamma)}{\gamma}\Big)^2 \gg \delta^2 \sum_{|\gamma|\leq 1/\delta} |M(1/2+i\gamma)|^2.$$

We may therefore focus on bounding J.

Note that

$$\frac{\xi'}{\xi}(s) = \Big(\frac{1}{s} + \frac{1}{s-1} - \log\sqrt{\pi} + \frac{1}{2}\frac{\Gamma'}{\Gamma}(s/2)\Big) + \frac{\zeta'}{\zeta}(s),$$

and accordingly write $J = J_1 + J_2$. To estimate J_2, we expand the Dirichlet series for $\zeta'/\zeta(s)$, $M(s)$ and $M(1-s)$ and exchange the summations and integration. Thus

$$J_2 = -\sum_{n=1}^{\infty} \Lambda(n) \sum_{\substack{m_1, m_2 \in \mathcal{S}(y) \\ \sqrt{x}/y^{1/3} \le m_1, m_2 \le \sqrt{x}/y^{\frac{1}{4}}}} \frac{1}{m_2}\frac{1}{2\pi i}\int_{(c)} \Big(\frac{m_2}{nm_1}\Big)^s \Big(\frac{e^{\delta(s-1/2)} - e^{-\delta(s-1/2)}}{(s-1/2)}\Big)^2 ds.$$

By shifting contours appropriately, we find that

$$\frac{1}{2\pi i}\int_{(c)} \xi^s \Big(\frac{e^{\delta(s-1/2)} - e^{-\delta(s-1/2)}}{(s-1/2)}\Big)^2 ds = \begin{cases} \sqrt{\xi}(2\delta - |\log \xi|) & \text{if } e^{-2\delta} \le \xi \le e^{2\delta} \\ 0 & \text{otherwise.} \end{cases}$$

Since m_1 and m_2 are below $\sqrt{x}/y^{\frac{1}{4}}$, if $nm_1 \ne m_2$ then m_2/nm_1 lies outside the interval $(e^{-2\delta}, e^{2\delta})$. Thus

$$J_2 = -2\delta \sum_{\substack{m_2 \in \mathcal{S}(y) \\ \sqrt{x}/y^{1/3} \le m_2 \le \sqrt{x}/y^{1/4}}} \frac{1}{m_2} \sum_{\substack{nm_1 = m_2 \\ \sqrt{x}/y^{1/3} \le m_1 \le \sqrt{x}/y^{1/4}}} \Lambda(n) < 0.$$

Thus $J \le J_1$, and we are reduced to estimating J_1.

To estimate J_1 we move the line of integration to the line $\mathrm{Re}(s) = 1/2$. We encounter a pole at $s = 1$ whose residue is $4M(0)M(1)(e^{\delta/2} - e^{-\delta/2})^2$. Using Stirling's formula we find that the remaining integral on the $\mathrm{Re}(s) = 1/2$ line is

$$\ll \int_{-\infty}^{\infty} |M(1/2+it)|^2 \log(2+|t|)\Big(\frac{\sin(\delta t)}{t}\Big)^2 dt.$$

Splitting this integral into the intervals $\mathcal{I}_j$ as above, and appealing to the mean value theorem for Dirichlet polynomials we conclude that this quantity is

$$\ll \delta^2 \log x \Big(\frac{\sqrt{x}}{y^{\frac{1}{4}}} + \frac{1}{\delta}\Big) M(1).$$

Since $M(0) \le \sqrt{x}/y^{\frac{1}{4}}$ we conclude that

$$J \le J_1 \ll \delta^2 \log x \Big(\frac{\sqrt{x}}{y^{\frac{1}{4}}} + \frac{1}{\delta}\Big) M(1).$$

This proves (5) for the region $\mathcal{I}_0$, and as noted before the other cases follow similarly. Our proof of the Theorem is now complete.

Acknowledgments. This note was inspired by an unpublished manuscript of Ganguly, Pal, and Sankaranarayanan. I offered them joint authorship of this note, but they declined and preferred that I write it up separately. I thank them for their comments and valuable correspondence. I also thank Carl Pomerance for drawing my attention to Hafner's work [6].

References

[1] A. Balog, *On the distribution of integers having no large prime factors*, Astérisque., **147-148** (1987), 27-31.

[2] E. Croot, *Smooth numbers in short intervals*, Int. J. Number Theory, **3** (2007), 159–169.

[3] K. Dickman, *On the frequency of numbers containing prime factors of a certain relative magnitude*, Ark. Mat. Astr. Fys., **22** (1930), 1-14.

[4] S. Ganguly and G. Pal, *Integers without large prime factors in short intervals: conditional results*, Proc. Ind. Acad. Sci,, to appear.

[5] A. Granville, *Smooth numbers: Computational Number Theory and Beyond*, Proc. MSRI Conf. Algorithmic Number Theory: Lattices, Number Fields, Curves and Cryptography, Berkeley, (2000), J. Buhler and P. Stevenhagen eds., Cambridge University Press, 1-56 .

[6] J. Hafner, *On smooth numbers in short intervals under the Riemann hypothesis*, preprint, (1993).

[7] G. Harman, *Short intervals containing numbers without large prime factors*, Math. Proc. Cambridege Philos. Soc., **109** no: 1 (1991), 1-5.

[8] G. Harman, *Integers without large prime factors in short intervals and arithmetic progressions*, Acta Arith. XCL. 3 (1999), 279–289.

[9] A. Hildebrand and G. Tenenbaum, *Integers without large prime factors*, J. Thor. Nombres. Bordeaux., **5** (1993) no: 2, 411-484.

[10] H. Iwaniec and E. Kowalski, *Analytic number theory.* Amer. Math. Soc. Colloquium Publications, **53**, (2004).

[11] H. W. Lenstra, Jr., *Factoring integers with elliptic curves*, Ann. of Math. (2) 126 (1987), no. 3, 649–673.

[12] K. Matomäki, *A note on smooth numbers in short intervals*, Int. J. Number Theory, to appear, 4 pages.

[13] C. Pomerance, *The role of smooth numbers in number-theoretic algorithms*, Proceedings of International Congress of Mathematicians., **1, 2** (Zurich, 1994), Birkhauser, Basel (1995), 411-422.

[14] K. Soundararajan, *An asymptotic expansion related to the Dickman function*, preprint, available as `arXiv:1005.3494`.

[15] T.Z. Xuan, *On smooth integers in short intervals under the Riemann Hypothesis*, Acta Arith., **88** (1999), 327-332.

Department of Mathematics,
Stanford University,
450 Serra Mall, Building 380,
Stanford, CA 94305-2125, USA.
ksound@math.stanford.edu