ON THE NUMBER OF INTEGERS $\leqq x$ WHOSE PRIME FACTORS DIVIDE n

Dedicated to Hans Rademacher on the occasion of his seventieth birthday

BY

N. G. DE BRUIJN

If n and x are positive integers, then we let $f(n, x)$ denote the number mentioned in the title, i.e., the number of integers m with $1 \leqq m \leqq x$, $m \mid n^{\infty}$. (The notation $m \mid n^{\infty}$ means that m divides some power of n, or in other words, that all prime factors of m divide n.)

P. Erdös conjectured (in a letter to the author, December 2, 1960) that the average $M(x) = x^{-1} \sum_{n=1}^{x} f(n, x)$ can be written as

$$M(x) = x^{-1} F(x) = \exp((\log x)^{1/2+\varepsilon(x)}), \qquad \text{where } \varepsilon(x) \to 0 \text{ for } x \to \infty.$$

We shall show in this paper (Theorem 2) that this is true. In fact we can get a much more precise result, viz. that $\log M(x)$ is asymptotically equivalent to $(8 \log x)^{1/2} (\log \log x)^{-1/2}$. Needless to say, this is still very far from an asymptotic formula for $M(x)$ itself.

The asymptotic formula for the logarithm of the average does not change if we replace $\sum_{n=1}^{x} f(n, x)$ by $\sum_{n=1}^{x} f(n, n)$, which is also considered in Theorem 2. This may give an idea of how rough our result still is.

We shall derive Theorem 2 from Theorem 1, which has some interest in itself. It deals with the partial sums of the series that results from the harmonic series if every denominator n is replaced by the product of the primes that divide it. This result will be obtained in a classical way: We build the corresponding Dirichlet series $f(\sigma)$ (see Lemma 2), we derive asymptotic information about $f(\sigma)$ if $\sigma \to 0$ (provided by Lemma 1), and we translate this information into information concerning the partial sums. This translation is achieved by a Tauberian theorem of Hardy and Ramanujan (see Lemma 3).

LEMMA 1. *Let h be a positive constant. If $\sigma > 0$, we define*

$$A_h(\sigma) = \int_{3/2}^{\infty} \{\log (1 + x^{-1}(x^{\sigma} - 1)^{-1})\} (\log x)^{-h} \, dx.$$

Then we have, if $\sigma \to 0$, $\sigma > 0$, and if h is fixed,

$$A_h(\sigma) = h^{-1} \sigma^{-1} (\log \sigma^{-1})^{-h} + O\{\sigma^{-1} (\log \sigma^{-1})^{-h-1} \log \log \sigma^{-1}\}.$$

Proof. Throughout this proof we abbreviate

$$(\log \sigma^{-1})^{-1} = \eta.$$

Received April 24, 1961.

We first integrate from $\frac{3}{2}$ to x_1, where

$$x_1 = \sigma^{-1}\eta^2,$$

which is $> \frac{3}{2}$ if σ is small enough. We have $0 < x_1^\sigma - 1 \leqq 2\sigma \log x_1$ provided that σ is small enough (notice that $\sigma \log x_1$ tends to zero if σ tends to zero). It follows that, if $\frac{3}{2} \leqq x \leqq x_1$,

$$0 \leqq x^{1+\sigma} - x \leqq x_1(x_1^\sigma - 1) \leqq 2\,\sigma x_1 \log x_1,$$

whence $0 \leqq x^{1+\sigma} - x \leqq \frac{1}{2}$ in that interval, provided that σ is small enough. We can now apply the inequality

$$1 + w^{-1} < w^{-2} \qquad (0 < w < \tfrac{1}{2}),$$

with $w = x(x^\sigma - 1)$. Remarking that $x^\sigma - 1 > \sigma \log x$ (since $\sigma \log x > 0$) and $x \log x > \frac{1}{2}$ ($x \geqq \frac{3}{2}$), we obtain $x(x^\sigma - 1) > \frac{1}{2}\,\sigma > \sigma^2$, whence

$$\int_{3/2}^{x_1} \{\log\,(1 + x^{-1}(x^\sigma - 1)^{-1})\}(\log x)^{-h}\,dx < \int_{3/2}^{x_1} \log \sigma^{-4}(\log x)^{-h}\,dx.$$

It follows that the contribution of the interval $2 \leqq x \leqq x_1$ to our integral is $O\{(\log \sigma^{-2})(\log x_1)^{-h}x_1\} = O(\sigma^{-1}\eta^{h+1})$.

For the remaining integral from x_1 to ∞ we shall derive an upper estimate and a lower estimate. For the upper estimate, we remark that

$$\{\log(1 + x^{-1}(x^\sigma - 1)^{-1})\} < x^{-1}(x^\sigma - 1)^{-1} < (x\sigma \log x)^{-1}$$

for all $x > 1$, whence

$$\int_{x_1}^{\infty} \{\log\,(1 + x^{-1}(x^\sigma - 1)^{-1})\}(\log x)^{-h}\,dx < \int_{x_1}^{\infty} x^{-1}\sigma^{-1}(\log x)^{-h-1}\,dx$$
$$= (h\sigma)^{-1}(\log x_1)^{-h} = (h\sigma)^{-1}\eta^h + O(\sigma^{-1}\eta^{h+1} \log \log \sigma^{-1}).$$

It follows that

$$A_h(\sigma) < (h\sigma)^{-1}\eta^h + O(\sigma^{-1}\eta^{h+1} \log \log \sigma^{-1}).$$

For our lower estimate we shall use

$$\int_{x_1}^{\infty} > \int_{x_2}^{x_3}, \qquad \text{where } x_2 = \sigma^{-1}, \quad x_3 = \exp\{(\log \sigma^{-1})^{(h+1)/h}\}.$$

If $x_2 \leqq x \leqq x_3$, we have

$$x(x^\sigma - 1) > x\sigma \log x \geqq x_2\,\sigma \log x_2 = \eta^{-1}.$$

Applying the inequality

$$v^{-1} \log(1 + v) \geqq 1 - \tfrac{1}{2}v \qquad (0 < v < 1),$$

with $v = (x(x^\sigma - 1))^{-1}$, we deduce that

$$\log\{1 + (x(x^\sigma - 1))^{-1}\} \geqq \{x(x^\sigma - 1)\}^{-1}(1 - \tfrac{1}{2}\eta) \qquad (x_2 \leqq x \leqq x_3),$$

provided that σ is small enough. Furthermore we have, if $x_2 \leqq x \leqq x_3$, that

$$\sigma \log x \leqq \sigma \log x_3 = O(\sigma\eta^{-(h+1)/h}) = o(\eta),$$

whence, if σ is small enough,

$$x^\sigma - 1 < (1 + \eta)\sigma \log x \qquad (x_2 \leqq x \leqq x_3).$$

It follows that

$$\int_{x_2}^{x_3} > \sigma^{-1}(1 - \tfrac{1}{2}\eta)(1 + \eta)^{-1} \int_{x_2}^{x_3} x^{-1}(\log x)^{-h-1}\,dx.$$

The integral on the right equals

$$h^{-1}(\log x_2)^{-h} - h^{-1}(\log x_3)^{-h} = h^{-1}(\eta^h - \eta^{h+1}).$$

It follows that $A_h(\sigma) > (h\sigma)^{-1}\eta^h - O(\sigma^{-1}\eta^{h+1})$, and this completes the proof of the lemma.

LEMMA 2. *Let $\alpha(n)$ denote the product of the different primes dividing n ($n = 1, 2, 3, \cdots$), and let $f(\sigma)$ denote the sum of the Dirichlet series*

$$f(\sigma) = \textstyle\sum_{n=1}^{\infty} (\alpha(n))^{-1} n^{-\sigma}.$$

This series converges if $\sigma > 0$, and we have the asymptotic equivalence

$$\log f(\sigma) \sim \sigma^{-1}(\log \sigma^{-1})^{-1} \qquad (\sigma \to 0).$$

Proof. The Dirichlet series has the product expansion

$$f(\sigma) = \textstyle\prod_p \{1 + p^{-1-\sigma} + p^{-1-2\sigma} + p^{-1-3\sigma} + \cdots\} = \prod_p \{1 + p^{-1}(p^\sigma - 1)^{-1}\},$$

where p runs through the primes. If σ is a fixed positive number, the factors of this product are, with at most a finite number of exceptions, less than the corresponding factors of the Euler product expansion for $\{\zeta(1 + \sigma)\}^2$ (where ζ is the Riemann zeta function). In fact we have

$$1 + p^{-1}(p^\sigma - 1)^{-1} < 1 + 2p^{-1-\sigma} < (1 - p^{-1-\sigma})^{-2}$$

as soon as $p^\sigma > 2$. This settles the matter of convergence.

It is a direct consequence of well-known facts in prime number theory that there exists a positive constant C such that

$$\int_{3/2}^{x} \{(\log t)^{-1} - C(\log t)^{-2}\}dt < \pi(x) < \int_{3/2}^{x} \{(\log t)^{-1} + C(\log t)^{-2}\}\,dt,$$

for all $x \geqq \frac{3}{2}$, where $\pi(x)$ stands for the number of primes $\leqq x$. Consequently, if $g(x)$ is a monotonically decreasing positive function with

$$\int_{3/2}^{\infty} g(x)\,(\log x)^{-1}\,dx < \infty,$$

we have

$$\left|\sum_p g(p) - \int_{3/2}^{\infty} g(x)\,(\log x)^{-1}\,dx\right| < C\int_{3/2}^{\infty} g(x)\,(\log x)^{-2}\,dx.$$

Applying this with

$$g(x) = \log\{1 + x^{-1}(x^{\sigma} - 1)^{-1}\},$$

we infer that, with the notation of Lemma 1,

$$|\log f(\sigma) - A_1(\sigma)| < CA_2(\sigma).$$

The asymptotic formula for $\log f(\sigma)$ now follows at once from that lemma.

LEMMA 3. *Let $a_n \geqq 0$ $(n = 1, 2, \cdots)$, assume that*

$$f(\sigma) = \sum_{n=1}^{\infty} a_n\, n^{-\sigma}$$

converges for all $\sigma > 0$, and that

$$\log f(\sigma) \sim \sigma^{-1}(\log \sigma^{-1})^{-1} \qquad (\sigma > 0,\quad \sigma \to 0).$$

Then we have

$$\log \sum_{n \leqq x} a_n \sim (8 \log x/\log\log x)^{1/2} \qquad (x \to \infty).$$

This is a special case of a Tauberian theorem given (for general Dirichlet series) by Hardy and Ramanujan [2]. (For further generalizations of that Tauberian theorem we refer to [1] and [3].)

Combining Lemmas 2 and 3, we obtain

THEOREM 1. *If $\alpha(n)$ represents the product of the different primes dividing n $(n = 1, 2, 3, \cdots)$, then we have*

$$\log\left\{\sum_{n\leqq x} (\alpha(n))^{-1}\right\} \sim (8 \log x)^{1/2}(\log\log x)^{-1/2} \qquad (x \to \infty).$$

THEOREM 2. *Let $f(n, x)$ be the number of positive integers $\leqq x$ which are products of powers of prime factors of n. We put*

$$F(x) = \sum_{n\leqq x} f(n, x), \qquad G(x) = \sum_{n\leqq x} f(n, n).$$

Then we have, as $x \to \infty$,

$$\log(x^{-1}F(x)) \sim \log(x^{-1}G(x)) \sim (8 \log x)^{1/2}(\log\log x)^{-1/2}.$$

Proof. Noticing that $k \mid n^{\infty}$ is equivalent to $n \equiv 0 \pmod{\alpha(k)}$, we obtain

$$F(x) = \sum_{n\leqq x}\sum_{k\leqq x,\, k\mid n^{\infty}} 1 = \sum_{k\leqq x}\sum_{n\leqq x,\, n\equiv 0 (\mathrm{mod}\, \alpha(k))} 1 = \sum_{k\leqq x} [x/\alpha(k)],$$

where $[z]$ denotes the largest integer $\leqq z$. And

$$\begin{aligned} G(x) &= \sum_{n\leqq x}\sum_{k\leqq n,\, k\mid n^{\infty}} 1 \\ &= \sum_{k\leqq x}\sum_{n\leqq x,\, n\geqq k,\, n\equiv 0(\mathrm{mod}\,\alpha(k))} 1 \\ &= \sum_{k\leqq x}\{[x/\alpha(k)] - [k/\alpha(k)] + 1\}. \end{aligned}$$

From these formulas we deduce

$$F(x) = x\sum_{k \leqq x} (\alpha(k))^{-1} + O(x),$$
$$G(x) = \sum_{k \leqq x} (x - k)(\alpha(k))^{-1} + O(x),$$

and for the latter sum we have

$$\tfrac{1}{2}x\sum_{k \leqq x/2} (\alpha(k))^{-1} \leqq \sum_{k \leqq x} (x - k)(\alpha(k))^{-1} \leqq x\sum_{k \leqq x} (\alpha(k))^{-1}.$$

The theorem now follows at once from the previous one.

The author is indebted to P. T. Bateman and E. E. Kohlbecker for several corrections.

References

1. N. G. de Bruijn, *Pairs of slowly oscillating functions occurring in asymptotic problems concerning the Laplace transform*, Nieuw Arch. Wisk. (3), vol. 7 (1959), pp. 20–26.
2. G. H. Hardy and S. Ramanujan, *Asymptotic formulae for the distribution of integers of various types*, Proc. London Math. Soc. (2), vol. 16 (1917), pp. 112–132.
3. E. E. Kohlbecker, *Weak asymptotic properties of partitions*, Trans. Amer. Math. Soc., vol. 88 (1958), pp. 346–365.

Technological University
Eindhoven, Netherlands