

# ON THE NUMBER OF UNCANCELLED ELEMENTS IN THE SIEVE OF ERATOSTHENES

BY

N. G. DE BRUIJN

(Communicated by Prof. W. VAN DER WOUDE at the meeting of April 29, 1950)

## 1. Introduction.

Let, for  $x > 0$ ,  $y \geq 2$ ,  $\Phi(x, y)$  denote the number of positive integers  $\leq x$  which have no prime factors  $< y$ . Information on  $\Phi(x, y)$  for large values of  $x$  and  $y$  can be obtained from several points of view.

A. First, for  $x$  very large with respect to  $y$  (roughly  $\log x > C y / \log^2 y$ ), the following elementary formula gives a satisfactory estimate. Putting  $\prod_{p < y} p = Q$ , we have LEGENDRE's formula

$$\Phi(x, y) = \sum_{d|Q} \mu(d) \left[ \frac{x}{d} \right]$$

and hence, for  $y \geq 2$ ,

$$(1.1) \quad \left| \Phi(x, y) - x \prod_{p < y} \left(1 - \frac{1}{p}\right) \right| \leq \sum_{d|Q} \left( \frac{x}{d} - \left[ \frac{x}{d} \right] \right) \leq \sum_{d|Q} 1 = 2^{\pi(y)} < 2^y.$$

B. On the other hand, if  $x$  is relatively small, the information comes from the prime number theorem. If  $y \leq x \leq y^2$ , the uncanceled elements in the sieve are exactly the primes in the interval  $y \leq p \leq x$ . Starting from here, A. BUCHSTAB<sup>1)</sup> derived estimates for  $\Phi(x, y)$  in the regions  $y^2 \leq x \leq y^3$ ,  $y^3 \leq x \leq y^4$ , ...,  $y^n \leq x \leq y^{n+1}$ , .... He gave, however, no estimates holding uniformly in  $n$ . In the present paper these will be achieved, owing to several improvements on BUCHSTAB's method.

BUCHSTAB's result was the following one. Let, for  $u \geq 1$ , the function  $\omega(u)$  be defined by

$$(1.2) \quad \begin{cases} \omega(u) = u^{-1}, & (1 \leq u \leq 2) \\ \frac{d}{du} \{u \omega(u)\} = \omega(u-1), & (u \geq 2) \end{cases}$$

---

<sup>1)</sup> Rec. Math [Mat. Sbornik], (2), **44**, 1239–1246 (1937). BUCHSTAB's work was partly duplicated by S. SELBERG (Norsk. Mat. Tidsskr. **26**, 79–84 (1944)). SELBERG also proved (cf. (1.16) below) that  $x^{-1} \Phi(x, y) \log y$  is uniformly bounded for  $x \geq y \geq 2$  (Norske Vid. Selsk. Forh., Trondhjem, **19** (2), 3–6 (1946)).

The present paper can be read independently from these publications.

where for  $u = 2$  the right-hand derivative has to be taken. Then for  $u > 1$ ,  $u$  fixed, BUCHSTAB proved:

$$(1.3) \quad \lim_{y \rightarrow \infty} \Phi(y^u, y) y^{-u} \log y = \omega(u).$$

It is not very difficult to derive from (1.2) that  $\lim_{u \rightarrow \infty} \omega(u)$  exists (see (1.10) below). Furthermore, since we have <sup>2)</sup>

$$(1.4) \quad \prod_{p < y} \left(1 - \frac{1}{p}\right) \sim e^{-\gamma} \log y,$$

where  $\gamma$  is EULER's constant, it can be expected that

$$(1.5) \quad \lim_{u \rightarrow \infty} \omega(u) = e^{-\gamma}.$$

Namely, if we put

$$(1.6) \quad \Phi(x, y) = x \prod_{p < y} \left(1 - \frac{1}{p}\right) \cdot \psi(x, y),$$

we obtain from (1.1) and (1.3), respectively

$$\lim_{u \rightarrow \infty} \psi(y^u, y) = 1, \quad \lim_{u \rightarrow \infty} \lim_{y \rightarrow \infty} \psi(y^u, y) = e^{\gamma} \lim_{u \rightarrow \infty} \omega(u).$$

Formula (1.5) can be established indeed. It will follow from the closer investigation of  $\psi(x, y)$  to be carried out in the sequel. A direct proof can also be given (section 4).

In section 2 we shall prove <sup>3)</sup>

$$(1.7) \quad |\psi(y^u, y) - e^{\gamma} \log y \int_1^u y^{t-u} \omega(t) dt| < CR(y) \quad (u \geq 1, y \geq 2).$$

Here  $\omega(t)$  is BUCHSTAB's function, defined by (1.2).  $R(y)$  is a positive function satisfying  $R(y) \downarrow 0$  for  $y \rightarrow \infty$ ,  $R(y) > y^{-1}$  and <sup>4)</sup>

$$(1.8) \quad \begin{cases} |\pi(y) - \text{li } y| < y R(y) / \log y & (y \geq 2) \\ \int_y^\infty |\pi(t) - \text{li } t| \cdot t^{-2} dt < R(y) & (y \geq 2). \end{cases}$$

Suitable functions  $R(y)$  are known from the theory of primes, for instance  $R(y) = C \exp(-C \log^{\kappa} x)$ , with  $\kappa = \frac{1}{2}$  <sup>5)</sup> or  $\kappa = \frac{4}{7} - \varepsilon$  <sup>6)</sup>, and  $\kappa = 1$  if the RIEMANN hypothesis is correct.

<sup>2)</sup> See A. E. INGHAM, The Distribution of Prime Numbers, London, p. 22 (1932).

<sup>3)</sup> The  $C$ 's are absolute constants, not necessarily the same at each occurrence.

<sup>4)</sup> As usual,  $\pi(y)$  denotes the number of primes  $\leq y$ ;  $\text{li } y$  denotes the logarithmic integral.

<sup>5)</sup> See INGHAM loc. cit. p. 65.

<sup>6)</sup> N. TCHUDAKOFF, C. R. Acad. Sci. URSS, N. s. 21, 421-422 (1938).

From (1. 7) we can deduce that

$$(1. 9) \quad |\psi(y^u, y) - 1| < C \Gamma^{-1}(u) + CR(y) \quad (u \geq 1, y \geq 2).$$

This follows from the behaviour of the function  $\omega(t)$ . In a previous paper<sup>7)</sup> we proved that  $\lim_{t \rightarrow \infty} \omega(t)$  exists, and, if we denote it by  $A$ , that

$$(1. 10) \quad \omega(u) = A + O\{\Gamma^{-1}(u+1)\}.$$

This can also be proved independently in a few lines. If we write (1. 2) in the form  $u \omega'(u) = -\omega(u) + \omega(u-1)$ , we infer that

$$\omega'(u) \leq u^{-1} \max_{u-1 \leq t \leq u} |\omega'(t)| \quad (u \geq 2).$$

It follows that  $\omega'(u)$  is bounded for  $u \geq 2$ . Denoting the upper bound of  $|\omega'(t)|$  for  $u \leq t < \infty$  by  $M(u)$ , we find  $M(u) \leq u^{-1} M(u-1)$  ( $u \geq 3$ ), and so  $M(u) \leq C \Gamma^{-1}(u+1)$ . Now (1. 10) easily follows.

The error-term in (1. 10) is certainly not the best possible. The right order is probably something like  $\exp(-u \log u - u \log \log u)$ .

It easily follows from (1. 10) that

$$|\log y \int_1^u y^{t-u} \omega(t) dt - A| < C \Gamma^{-1}(u) + Cy^{-1} \quad (u \geq 1, y \geq 2).$$

Now (1. 7) leads to

$$(1. 11) \quad |\psi(y^u, y) - e^\gamma A| < C \Gamma^{-1}(u) + CR(y) \quad (u \geq 1, y \geq 2).$$

Take a fixed value of  $y$ , and make  $u \rightarrow \infty$ . Comparing the result with (1. 1) we find  $|e^\gamma A - 1| < CR(y)$  ( $y \geq 2$ ), and hence  $A = e^{-\gamma}$ . This proves (1. 5) and (1. 9).

C. There is a third approach to the problem of  $\Phi(x, y)$ . Put

$$\zeta(s) \prod_{p < y} (1 - p^{-s}) = \sum_{n=1}^{\infty} c_n n^{-s}, \quad (\text{Re } s > 1),$$

then

$$(1. 12) \quad \Phi(x, y) = \sum_{n \leq x} c_n,$$

and this sum can be evaluated by contour integration. This will be exposed in section 3. There is nothing new in the method which is quite familiar from the theory of DIRICHLET series. The result has some interest, since, for  $u = (\log x)/\log y$  not too large, it does not fall far from (1. 9). We shall prove in section 3, namely, that

$$(1. 13) \quad \begin{cases} |\psi(y^u, y) - 1| < C \log^3 y \cdot e^{-u \log u - u \log \log u + Cu} \\ (1 \leq u \leq 4y^{1/\log y}; \quad y \geq 2). \end{cases}$$

<sup>7)</sup> N. G. DE BRUIJN, On some linear functional equations, Example 1. To be published in *Publicationes Mathematicae*, Debrecen.

On the other hand we shall show that

$$(1.14) \quad |\psi(y^u, y) - 1| < C e^{-iu \log y} = C x^{-i} \quad (u > 4y^{1/2}/\log y; y \geq 2).$$

For  $u > \varepsilon^{-1} y \log^{-2} y$ , however, (1.1) gives a better result than this one, namely

$$(1.15) \quad |\psi(x, y) - 1| < C x^{-1+\varepsilon} \quad (u > \varepsilon^{-1} y \log^{-2} y; y \geq 2).$$

In section 3 we have to make the restrictions  $y \geq e^2$ ,  $u > 2e$ , but it is easily seen from (1.9) and (1.15) that in (1.13) and (1.14) these restrictions may be removed.

It is easily inferred from the results of methods **B** and **C** that there is a positive constant  $\alpha$ , such that

$$(1.16) \quad |\psi(y^u, y) - 1| < C e^{-\alpha u} \quad (y \geq 2, u \geq 1).$$

Namely, in (1.9) we can take  $R(y) = C \exp(-C \log^3 y)$ . Hence (1.16) holds, with  $\alpha = 1$ , for  $y \geq 2$ ,  $1 \leq u \leq C \log^3 y$ . On the other hand, if  $u > C \log^3 y$ , we have  $\log^3 y < C u^6 < C e^u$ . Consequently, by (1.13) and (1.14), we have

$$|\psi(y^u, y) - 1| < C \text{Max} (e^{-u \log u - u \log \log u + Cu}, e^{-iu \log y}).$$

This proves (1.16), with  $\alpha = \frac{1}{2} \log 2$ .

BUCHSTAB considered, in the paper quoted before, the more general problem of the uncanceled numbers which belong to a given arithmetical progression. Suppose  $k \geq 1$ ,  $(l, k) = 1$ , and let  $\Phi_l(k; x, y)$  denote the number of positive integers  $\leq x$ , which are  $\equiv 1 \pmod{k}$ , and which contain no prime factors  $< y$ . For  $\varphi(k) \Phi_l(k; x, y)$ , where  $\varphi(k)$  is EULER's indicator, he obtained the same result as for the case  $k = 1$  described above. The present results can also be generalised that way. In section 2 this can be carried out with very little alterations. Following BUCHSTAB, we can simply deal simultaneously with all  $\Phi_l(k; x, y)$ , for  $k, x, y$  fixed. In section 3 we have to use DIRICHLET's  $L$ -series instead of the RIEMANN zeta function. In both methods real difficulties only arise when estimations holding uniformly in  $k$  are required.

## 2. Proof of (1.7).

Suppose  $x \geq y \geq 2$ . Clearly, for  $h \geq 1$

$$(2.1) \quad \Phi(x, y) = \sum_{y \leq p < y^h} \Phi\left(\frac{x}{p}, p\right) + \Phi(x, y^h),$$

and hence, by (1.6)

$$(2.2) \quad \psi(x, y) = \sum_{y \leq p < y^h} \psi\left(\frac{x}{p}, p\right) \cdot \frac{1}{p} \prod_{y \leq q < p} \left(1 - \frac{1}{q}\right) + \psi(x, y^h) \prod_{y \leq p < y^h} \left(1 - \frac{1}{p}\right).$$

Here  $p$  and  $q$  run through the primes. Put

$$\prod_{p < y} \left(1 - \frac{1}{p}\right) = P(y), \quad \sum_{y \leq p < y^\sigma} \frac{1}{p} \prod_{y \leq q < p} \left(1 - \frac{1}{q}\right) = W(\sigma) = 1 - \frac{P(y^\sigma)}{P(y)}.$$

$W(\sigma)$  depends on  $y$  also.

Using a STIELTJES integral, we can write instead of (2. 2), for  $u \geq 1$ ,

$$(2. 3) \quad \psi(y^u, y) = \int_1^h \psi(y^{u-\sigma}, y^\sigma) dW(\sigma) + \psi(y^u, y^h) \{1 - W(h)\}.$$

We first estimate  $W(\sigma)$ . We have for  $\sigma \geq 1$ , if  $\pi^*(y)$  denotes the number of primes  $< y$  (not  $\leq y$ ):

$$\begin{aligned} \log \{P(y^\sigma)/P(y)\} &= \sum_{y \leq p < y^\sigma} \log \left(1 - \frac{1}{p}\right) = \int_1^\sigma \log(1 - y^{-\mu}) d\pi^*(y^\mu) = \\ &= \int_1^\sigma \log(1 - y^{-\mu}) d\{\pi^*(y^\mu) - \text{li } y^\mu\} + \int_1^\sigma \log(1 - y^{-\mu}) \frac{y^\mu \log y}{\log y^\mu} d\mu = \\ &= \log(1 - y^{-\mu}) \{\pi^*(y^\mu) - \text{li } y^\mu\} \Big|_1^\sigma - \int_1^\sigma \frac{\log y}{y^\mu - 1} \{\pi^*(y^\mu) - \text{li } y^\mu\} d\mu - \log \sigma + O\left(\frac{1}{y}\right). \end{aligned}$$

By (1. 8) we now find

$$(2. 4) \quad |\sigma P(y^\sigma)/P(y) - 1| < C R(y) \quad (\sigma \geq 1, y \geq 2)$$

Hence we have

$$(2. 5) \quad |W(\sigma) - 1 + \sigma^{-1}| < C R(y). \quad (\sigma \geq 1, y \geq 2)$$

An approximate solution of (2. 3) is  $\theta(y^u, y)$ , where

$$(2. 6) \quad \theta(y^u, y) = e^\gamma \log y \cdot \int_1^h y^{t-u} \omega(t) dt,$$

and  $\omega(t)$  is given by (1. 2). In order to show this, we first evaluate

$$\Omega_1(u; y; h) = \theta(y^u, y) - \int_1^h \theta(y^{u-\sigma}, y^\sigma) \sigma^{-2} d\sigma - h^{-1} \theta(y^u, y^h)$$

for  $1 \leq h \leq \frac{1}{2} u$ . We have  $\Omega(u; y; 1) = 0$ , and

$$\frac{\partial}{\partial h} \Omega_1(u; y; h) = -h^{-2} \theta(y^{u-h}, y^h) + h^{-2} \theta(y^u, y^h) - h^{-1} \frac{\partial}{\partial h} \theta(y^u, y^h).$$

The right-hand-side can be evaluated; using (2. 6) and (1. 2) we find it to be  $-e^\gamma h^{-1} y^{h-u} \log y$ . Therefore

$$(2. 7) \quad |\Omega_1(u; y; h)| = e^\gamma \log y \int_1^h t^{-1} y^{t-u} dt \leq e^\gamma y^{h-u}.$$

Next we put

$$\begin{aligned} \Omega_3(u; y; h) &= \theta(y^u, y) - \int_1^h \theta(y^{u-\sigma}, y^\sigma) dW(\sigma) - \theta(y^u, y^h) \{1 - W(h)\} = \\ &= \Omega_1(u; y; h) + \Omega_2(u; y; h) \end{aligned}$$

The most important contribution is  $J_1$ . We have

$$(3.3) \quad \int_{b-iT}^{b+iT} \left| \zeta(s) \frac{ds}{s} \right| = O(\log^{3/2} T) + O\left(\log \frac{1}{1-b}\right).$$

This is easily deduced from the fact that, for  $T_1 > 2$ ,  $\frac{1}{2} \leq b \leq 1$ ,

$$\int_{T_1}^{2T_1} |\zeta(b \pm it)|^2 dt = O(T_1 \log T_1)^8$$

whereas the second  $O$ -term in (3.3) arises from the pole of  $\zeta(s)$  at  $s = 1$ .

Furthermore, for  $s = b + it$  we simply use

$$\left| \prod_{p < y} (1 - p^{-s}) \right| \leq \prod_{p < y} (1 + p^{-b}),$$

and, as  $\pi(\xi) < 2 \operatorname{li} \xi + C$ ,

$$\begin{aligned} \log \prod_{p < y} (1 + p^{-b}) &\leq O(1) + \int_2^y \xi^{-b} d\pi(\xi) \leq O(1) + 2 \int_2^y \xi^{-b} d \operatorname{li} \xi = \\ &= O(1) + 2 \int_{(1-b) \log 2}^{\lambda} e^{\eta} \eta^{-1} d\eta = O(1) + 2 \operatorname{li} e^{\lambda} + 2 \log \log y - 2 \log \lambda. \end{aligned}$$

Summarizing, we have

$$(3.4) \quad J_1 = O \left\{ x e^{-\lambda u} \cdot 2 \operatorname{li} e^{\lambda} \cdot \frac{(\log y)^2}{\lambda^2} \left( \log^{3/2} T + \log \frac{\log y}{\lambda} \right) \right\}.$$

For  $J_2$  we use

$$|\zeta(\sigma + it)| < C t^{\frac{1}{2}}, \quad (|t| \geq 2, \sigma \geq \tfrac{1}{2})$$

which leads to

$$(3.5) \quad J_2 = O \left\{ x^a \exp(2 \operatorname{li} e^{\lambda}) \cdot \left( \frac{\log y}{\lambda} \right)^2 \cdot T^{-1} \right\}.$$

The same holds for  $J_3$ , of course.

The difference  $\Phi(x, y) - x \prod_{p < y} (1 - p^{-1})$  is less than the sum of (3.2),

(3.4) and (3.5). Simplifying our result by specialization

$$a = 1 + (\log x)^{-1}, \quad T = e^{2\lambda u},$$

we can deduce

$$\Phi(x, y) - x \prod_{p < y} (1 - p^{-1}) = O \{ x \log^2 y \cdot (\lambda u)^{3/2} \cdot \exp(-\lambda u + 2 \operatorname{li} e^{\lambda}) \}.$$

By (1.6) and (1.4) we now obtain

$$(3.6) \quad \psi(x, y) - 1 = O \{ \log^3 y \cdot (\lambda u)^{3/2} \cdot \exp(-\lambda u + 2 \operatorname{li} e^{\lambda}) \}.$$

We are still free to give  $\lambda$  any value in the interval  $1 \leq \lambda \leq \frac{1}{2} \log y$ .

<sup>8)</sup> TITCHMARSH, The Zeta Function of RIEMANN. London, p. 31 (1931).

Now assume that  $u > 2e$ . Then the minimum of  $-\lambda u + 2 \operatorname{li} e^\lambda$  for  $\lambda \geq 1$  is attained for  $\lambda = \lambda_0$ , defined by

$$(3.7) \quad \lambda_0 u = 2e^{\lambda_0}, \quad \lambda_0 > 1.$$

There we have (put  $2\xi^{-1}e^\xi = \eta$ )

$$\operatorname{li} e^{\lambda_0} = O(1) + \int_1^{\lambda_0} e^\xi \xi^{-1} d\xi = O(1) + \frac{1}{2} \int_{2e}^u (1 - \xi^{-1})^{-1} d\eta = O(u).$$

Furthermore

$$\lambda_0 = \log(\tfrac{1}{2}u\lambda_0) > \log \tfrac{1}{2}u + \log \log \tfrac{1}{2}u > \log u + \log \log u + O(1),$$

whence it follows

$$(3.8) \quad \exp(-\lambda_0 u + 2 \operatorname{li} e^{\lambda_0}) < \exp\{-u \log u - u \log \log u + O(u)\}.$$

This result can be applied to (3.6) whenever the solution of (3.7) is less than  $\frac{1}{2} \log y$ , that is for

$$2e < u \leq 4y^{\frac{1}{2}}/\log y.$$

In that region we obtain, since  $\lambda_0 u = O(e^u)$ ,

$$\psi(x, y) - 1 = O(\log^3 y \cdot e^{-u \log u - u \log \log u + Cu}).$$

If, however,  $u > 4y^{\frac{1}{2}}/\log y$ , we take  $\lambda = \frac{1}{2} \log y$ , and we infer from (3.6)

$$\psi(x, y) - 1 = O(e^{-\frac{1}{2}u \log y}) = O(x^{-1/2}).$$

#### 4. Direct proof of (1.4).

We have established by combination of the results of methods A and B (cf. section 1), that

$$(4.1) \quad \lim_{u \rightarrow \infty} \omega(u) = e^{-\gamma}.$$

A purely analytical proof can also be given. The function

$$(4.2) \quad h(u) = \int_0^\infty \exp\{-ux - x + \int_0^x (e^{-t} - 1)t^{-1} dt\} dx$$

is analytical for  $u > -1$ ; it satisfies the equation

$$(4.3) \quad u h'(u-1) + h(u) = 0 \quad (u > 0).$$

The expression

$$\int_{a-1}^a \omega(u) h(u) du + a \omega(a) h(a-1) = (\omega, h)$$

does not depend on  $a$  for  $a \geq 2$ . This is easily verified by differentiation, using (1.2) and (4.3). We now evaluate  $(\omega, h)$  in two ways. First let  $a \rightarrow \infty$ . Then we have  $\omega(a) \rightarrow A$ , by (1.10), and  $h(a) \sim a^{-1}$  by (4.2). Hence  $(\omega, h) = A$ . Secondly, take  $a = 2$ . We find

$$A = \int_1^2 u^{-1} h(u) du + h(1).$$

In virtue of (4.3) we obtain

$$\begin{aligned} A &= - \int_1^2 h'(u-1) du + h(1) = h(0) = \lim_{u \downarrow 0} u h'(u-1) = \\ &= - \lim_{u \downarrow 0} u \int_0^\infty \exp \left\{ -ux + \int_0^\infty (e^{-t}-1) t^{-1} dt + \log x \right\} dx. \end{aligned}$$

Here we have

$$\lim_{x \rightarrow \infty} \left\{ \int_0^\infty \frac{e^{-t}-1}{t} dt + \log x \right\} = -\gamma,$$

and (4.1) easily follows.

The method used here consists of the construction of an “adjoint” equation (4.3), such that there is an invariant inner product  $(\omega, h)$  for any pair of solutions  $\omega, h$  of the original equation and of the adjoint one, respectively. This method was decisive in the author’s researches on the equation  $F'(x) = e^{ax+\beta} F(x-1)$ , which are unpublished as yet.

*Mathematisch Instituut  
der Technische Hogeschool,  
Delft, Netherlands.*