# AN ELEMENTARY PROOF OF THE PRIME NUMBER THEOREM WITH REMAINDER TERM

ROBERT BREUSCH

**Introduction.** In this paper, the prime number theorem in the form $\psi(x) \equiv \sum_{p^m \leqq x} \log p = x + o(x \cdot \log^{-1/6+\varepsilon} x)$, for every $\varepsilon > 0$, is established via a proof that in the well-known formula

$$(1) \qquad \rho(x) \equiv \sum_{p^m \leqq x} \frac{\log p}{p^m} = \log x + O(1) \equiv \log x + a_x ,$$

$a_x = -A_0 + o(\log^{-1/6+\varepsilon} x)$. ($A_0$ is Euler's constant.)

Throughout the paper, $p$ and $q$ stand for prime numbers, $k, m, n, t$, and others are positive integers, and $x, y$, and $z$ are positive real numbers.

Some well-known formulas, used in the proof, are

$$(2) \qquad \sum_{n \leqq x} \frac{\log^k n}{n} = \frac{1}{k+1} \cdot \log^{k+1} x + A_k + O\left(\frac{\log^k x}{x}\right), \quad \text{for } k = 0, 1, \cdots$$

$$(2') \qquad \sum_{y < n \leqq z} \frac{1}{n} \cdot \log^k(n/y) = \frac{1}{k+1} \cdot \log^{k+1}(z/y) + O\left(\frac{1}{y} \cdot \log^k(z/y)\right), \quad \text{for } k = 0, 1, \cdots$$

$$(3) \qquad \sum_{n \leqq x} \log^k(x/n) = O(x), \quad \text{for } k = 1, 2, \cdots$$

$$(4) \qquad \sum_{p^m \leqq x} \log p \cdot \log^k(x/p^m) = O(x), \quad \text{for } k = 0, 1, \cdots$$

$$(5) \qquad \sum_{n \leqq x} \mu(n)/n = O(1) \quad (\mu(n) \text{ is Moebius' function.})$$

Two other formulas, used prominently, are

$$(6) \qquad \sigma(x) \equiv \sum_{p^m \leqq x} \frac{\log p}{p^m} \cdot \log(x/p^m) = \frac{1}{2} \cdot \log^2 x - A_0 \cdot \log x + g_x \ (g_x = O(1))$$

$$(7) \qquad \tau(x) \equiv \sum_{p^m \leqq x} \frac{\log p}{p^m} \cdot \log^2(x/p^m) = \frac{1}{3} \cdot \log^3 x - A_0 \cdot \log^2 x + (2 \cdot A_0^2 + 4 \cdot A_1) \log x + O(1) .$$

With the help of (1), (2), and (4), (6) can be proved easily:

$$\sigma(x) = \sum_{p^m \leqq x} \frac{\log p}{p^m} \cdot \Bigl( \sum_{n \leqq x/p^m} 1/n - A_0 + O(p^m/x) \Bigr), \quad \text{or, with } k = n \cdot p^m ,$$

$$\sigma(x) = \sum_{k \leqq x} \frac{1}{k} \cdot \sum_{p^m / k} \log p - A_0 \cdot \log x + O(1)$$

$$= \sum_{k \leqq x} \frac{\log k}{k} - A_0 \log x + O(1) = \frac{1}{2} \cdot \log^2 x - A_0 \log x + O(1) .$$

Received December 31, 1958.

Also, again with $k = n \cdot p^m$,

$$\sum_{k\leqq x} \frac{\log k}{k} \cdot \log \frac{x}{k}$$
$$= \sum_{k\leqq x} \frac{1}{k} \cdot \log \frac{x}{k} \cdot \sum_{p^m/k} \log p = \sum_{p^m\leqq x} \log p \cdot \sum_{n\leqq x/p^m} \frac{1}{n\cdot p^m} \cdot \log\left(\frac{x}{n\cdot p^m}\right)$$
$$= \sum_{p^m\leqq x} \frac{\log p}{p^m} \cdot \left\{\log\left(\frac{x}{p^m}\right) \cdot \sum_{n\leqq x/p^m} \frac{1}{n} - \sum_{n\leqq x/p^m} \frac{\log n}{n}\right\}$$
$$= \sum_{p^m\leqq x} \frac{\log p}{p^m} \cdot \left\{\log^2 (x/p^m) + A_0 \log (x/p^m) - \frac{1}{2} \log^2 (x/p^m) - A_1\right\} + O(1)$$

(by (2) and (4))

$$= \frac{1}{2} \cdot \tau(x) + A_0 \cdot \sigma(x) - A_1 \cdot \rho(x) + O(1) .$$

(7) follows now by (1), (2), and (6).

The proof now proceeds in the following steps: in part I, certain asymptotic formulas for $a_n$ (see (1)) and $g_n$ (see (6)) are derived; they suggest that "on the average," $a_n$ is $-A_0$, and $g_n$ is $A_0^2 + 2A_1$. In part II, formulas for $a_n$ and $g_n$ are derived which are of the type of Selberg's asymptotic formula for $\psi(x)$; part III contains the final proof.

## PART I

First, the following five formulas will be derived; $K_1, K_2, \cdots,$ are constants, independent of $x$.

$$(8) \qquad \sum_{n\leqq x} \frac{1}{n} \cdot a_n = -A_0 \log x + g_x + K_2 + O\left(\frac{\log x}{x}\right)$$

$$(9) \qquad \sum_{n\leqq x} \frac{1}{n} \cdot a_{x/n} = -A_0 \log x + K_3 + O\left(\frac{\log x}{x}\right)$$

$$(10) \qquad \sum_{p\ \leqq x} \frac{\log p}{p^m} \cdot a_{p^m} = -A_0 \log x + g_x + \frac{1}{2} a_x^2 + K_4 + O\left(\frac{\log x}{x}\right)$$

$$(11) \qquad \sum_{n\leqq x} \frac{1}{n} \cdot g_n = (A_0^2 + 2 \cdot A_1) \cdot \log x + O(1)$$

$$(12) \qquad \sum_{n\leqq x} \frac{1}{n} \cdot g_{x/n} = (A_0^2 + 2 \cdot A_1) \cdot \log x + K_5 + O\left(\frac{\log^2 x}{x}\right) .$$

*Proofs.*

$$\sigma(x) = \sum_{n\leqq x} \log \frac{x}{n} (\rho(n) - \rho(n-1)) = \sum_{n\leqq x} \rho(n) \cdot \log \frac{n+1}{n} + O\left(\frac{\log x}{x}\right)$$
$$= \sum_{n\leqq x} \frac{\rho(n)}{n} + K_1 + O\left(\frac{\log x}{x}\right)$$
$$= \sum_{n\leqq x} \frac{\log n}{x} + \sum_{n\leqq x} \frac{1}{n} a_n + K_1 + O\left(\frac{\log x}{x}\right) .$$

(8) follows now from (6) and (2).

Also

$$\sum_{n\leqq x}\frac{1}{n}\cdot a_{x/n}=\sum_{n\leqq x}\frac{1}{n}\cdot\Big(\sum_{p^m\leqq x/n}\frac{\log p}{p^m}-\log\frac{x}{n}\Big)$$
$$=\sum_{k\leqq x}\frac{1}{k}\sum_{p^m/k}\log p-\sum_{n\leqq x}\frac{1}{n}\log\frac{x}{n}\quad(k=n\cdot p^m)$$
$$=\sum_{k\leqq x}\frac{\log k}{k}-\sum_{n\leqq x}\frac{1}{n}\log\frac{x}{n}\,.\quad\text{which proves (9) by (2).}$$

And

$$\sum_{p^m\leqq x}\frac{\log p}{p^m}\cdot a_{p^m}=\sum_{p^m\leqq x}\frac{\log p}{p^m}\Big(\sum_{q^t\leqq p^m}\frac{\log q}{q^t}-\log(p^m)\Big)$$
$$=\frac{1}{2}\Big(\sum_{p^m\leqq x}\frac{\log p}{p^m}\Big)^2+\frac{1}{2}\sum_{p^m\leqq x}\frac{\log^2 p}{p^{2m}}-\log x\cdot\sum_{p^m\leqq x}\frac{\log p}{p^m}+\sum_{p^m\leqq x}\frac{\log p}{p^m}\log\frac{x}{p^m}\,.$$

Thus, by (1), (2) and (6),

$$\sum_{p^m\leqq x}\frac{\log p}{p^m}\cdot a_{p^m}=\frac{1}{2}(\log x+a_x)^2+K_4+O\Big(\frac{\log x}{x}\Big)-\log x\cdot(\log x+a_x)$$
$$+\frac{1}{2}\log^2 x-A_0\log x+g_x\,,\quad\text{which proves (10).}$$

In the next proof, use is made of the easily established fact that

$$\rho(n)\cdot\log\frac{n+1}{n}=\sigma(n+1)-\sigma(n)\,.$$

$$\tau(x)=\sum_{n\leqq x}\log^2\Big(\frac{x}{n}\Big)(\rho(n)-\rho(n-1))$$
$$=\sum_{n\leqq x}\rho(n)\Big(\log^2\Big(\frac{x}{n}\Big)-\log^2\Big(\frac{x}{n+1}\Big)\Big)+O(1)$$
$$=\sum_{n\leqq x}\rho(n)\log\frac{n+1}{n}\cdot\log\frac{x^2}{n(n+1)}+O(1)$$
$$=\sum_{n\leqq x}(\sigma(n+1)-\sigma(n))\cdot\log\frac{x^2}{n(n+1)}+O(1)$$
$$=\sum_{n\leqq x}\sigma(n)\cdot\log\frac{n+1}{n-1}+O(1)=\sum_{n\leqq x}\sigma(n)\cdot\frac{2}{n}+O(1)$$
$$=\sum_{n\leqq x}\frac{\log^2 n}{n}-2\cdot A_0\cdot\sum_{n\leqq x}\frac{\log n}{n}+2\cdot\sum_{n\leqq x}\frac{1}{n}\cdot g_n+O(1)\quad\text{(by (6)).}$$

This proves (11), with the help of (2) and (7).

Finally

$$\sum_{n\leqq x}\frac{1}{n}\cdot g_{x/n}=\sum_{n\leqq x}\frac{1}{n}\cdot\Big(\sum_{p^m\leqq x/n}\frac{\log p}{p^m}\cdot\log\frac{x}{n\cdot p^m}-\frac{1}{2}\log^2\Big(\frac{x}{n}\Big)+A_0\log\frac{x}{n}\Big)\,,$$

or, with $k=n\cdot p^m$,

$$\sum_{n\leq x}\frac{1}{n}\cdot g_{x/n}$$
$$=\sum_{k\leq x}\frac{1}{k}\cdot\log\frac{x}{k}\cdot\sum_{p^m/k}\log p-\frac{1}{2}\cdot\sum_{n\leq x}\frac{1}{n}\cdot\log^2\left(\frac{x}{n}\right)+A_0\sum_{n\leq x}\frac{1}{n}\cdot\log\frac{x}{n}$$
$$=\sum_{k\leq x}\frac{1}{k}\cdot\log\frac{x}{k}\cdot\log k-\frac{1}{2}\cdot\sum_{n\leq x}\frac{1}{n}\cdot\log^2\left(\frac{x}{n}\right)+A_0\sum_{n\leq x}\frac{1}{n}\cdot\log\frac{x}{n}\,.$$

(12) now follows by (2).

Formulas (8) through (12) suggest setting

$$b_x\equiv a_x+A_0\,,\qquad h_x\equiv g_x-(A_0^2+2A_1)\,. \tag{13}$$

In terms of $b_x$ and $h_x$, the five formulas read

$$\sum_{n\leq x}\frac{1}{n}\cdot b_n=h_x+K_6+O\left(\frac{\log x}{x}\right) \tag{8'}$$

$$\sum_{n\leq x}\frac{1}{n}\cdot b_{x/n}=K_7+O\left(\frac{\log x}{x}\right) \tag{9'}$$

$$\sum_{p^m\leq x}\frac{\log p}{p^m}\cdot b_{p^m}=A_0\cdot(-A_0+b_x)+A_0^2+2\cdot A_1+h_x+K_4+\frac{1}{2}\cdot(-A_0+b_x)^2+O\left(\frac{\log x}{x}\right)$$
$$=h_x+\frac{1}{2}\cdot b_x^2+K_8+O\left(\frac{\log x}{x}\right) \tag{10'}$$

$$\sum_{n\leq x}\frac{1}{n}\cdot h_n=O(1) \tag{11'}$$

$$\sum_{n\leq x}\frac{1}{n}\cdot h_{x/n}=K_9+O\left(\frac{\log^2 x}{x}\right)\,. \tag{12'}$$

Next, it will be shown that

$$\sum_{n\leq x}\frac{1}{n}\cdot b_n^2=\sum_{n\leq x}\frac{1}{n}\cdot b_{x/n}^2+O(1)\,, \tag{14}$$

and

$$\sum_{n\leq x}\frac{1}{n}\cdot b_n\cdot b_{x/n}=\sum_{p^m\leq x}\frac{\log p}{p^m}\cdot h_{x/p^m}+O(1)\,. \tag{15}$$

For a proof of (14), we know, by (10′), that

$$\frac{1}{n}\cdot b_n^2=\frac{2}{n}\cdot\sum_{p^m\leq n}\frac{\log p}{p^m}\cdot b_{p^m}-\frac{2}{n}\cdot h_n-\frac{2}{n}\cdot K_8+O\left(\frac{\log n}{n^2}\right)\,,$$

and

$$\frac{1}{n}\cdot b_{x/n}^2=\frac{2}{n}\cdot\sum_{p^m\leq x/n}\frac{\log p}{p^m}\cdot b_{p^m}-\frac{2}{n}\cdot h_{x/n}-\frac{2}{n}\cdot K_8+O\left(\frac{1}{x}\cdot\log\frac{x}{n}\right)\,.$$

Thus, by (3), (11′) and (12′),

$$\sum_{n\le x}\frac{1}{n}\cdot(b_n^2-b_{x/n}^2)=2\cdot\sum_{n\le x}\frac{1}{n}\Big(\sum_{p^m\le n}\frac{\log p}{p^m}\cdot b_{p^m}-\sum_{p^m\le x/n}\frac{\log p}{p^m}\cdot b_{p^m}\Big)+O(1)$$

$$=2\cdot\sum_{p^m\le x}\frac{\log p}{p^m}\cdot b_{p^m}\Big(\sum_{p^m\le n\le x}\frac{1}{n}-\sum_{n\le x/p^m}\frac{1}{n}\Big)+O(1)$$

$$=2\cdot\sum_{p^m\le x}\frac{\log p}{p^m}\cdot b_{p^m}\cdot\Big(\log(x/p^m)+O(1/p^m)-\log(x/p^m)$$

$$-A_0-O(p^m/x)\Big)+O(1)$$

$=O(1)$, by (10′) and (4). This proves (14).

Also

$$\sum_{n\le x}\frac{1}{n}\cdot b_n\cdot b_{x/n}=\sum_{n\le x}\frac{1}{n}\cdot b_n\cdot\Big(\sum_{p^m\le x/n}\frac{\log p}{p^m}-\log\frac{x}{n}+A_0\Big)$$

$$=\sum_{n\le x}\frac{1}{n}\cdot b_n\cdot\Big(\sum_{p^m\le x/n}\frac{\log p}{p^m}-\sum_{t\le x/n}\frac{1}{t}+2\cdot A_0+O\Big(\frac{n}{x}\Big)\Big)$$

$$=\sum_{p^m\le x}\frac{\log p}{p^m}\sum_{n\le x/p^m}\frac{1}{n}b_n-\sum_{t\le x}\frac{1}{t}\sum_{n\le x/t}\frac{1}{n}b_n+O(1)\,,\quad\text{by (8′)}$$

$$=\sum_{p^m\le x}\frac{\log p}{p^m}\cdot h_{x/p^m}+K_6\log x-\sum_{t\le x}\frac{1}{t}h_{x/t}-K_6\log x+O(1)$$

$$\text{(by (8′), (1) and (4))}$$

$$=\sum_{p^m\le x}\frac{\log p}{p^m}\cdot h_{x/p^m}+O(1)\,,\quad\text{by (12′)}\,.$$

From (14) and (15) it follows that

$$\sum_{n\le x}\frac{1}{n}\cdot(b_n\pm b_{x/n})^2=2\cdot\sum_{n\le x}\frac{1}{n}\cdot b_n^2\pm2\cdot\sum_{p^m\le x}\frac{\log p}{p^m}\cdot h_{x/p^m}+O(1)\,,$$

and therefore

$$\sum_{n\le x}\frac{1}{n}\cdot b_n^2\ge\Big|\sum_{p^m\le x}\frac{\log p}{p^m}\cdot h_{x/p^m}\Big|+O(1)\,. \tag{16}$$

## PART II

In the following, we shall employ the inversion formula

$$G(x)=\sum_{n\le x}g\Big(\frac{x}{n}\Big)\qquad\text{for all } x>0\Rightarrow g(x)=\sum_{n\le x}\mu(n)\cdot G\Big(\frac{x}{n}\Big)\,,$$

as well as

$$\sum_{n\le x}\frac{\mu(n)}{n}\log\frac{x}{n}=O(1)\,. \tag{17}$$

For a proof of (17), we make use of the fact that $\sum_{n\le x}x/n=x\cdot\log x+A_0x+O(1)$; thus, by the inversion formula,

$$x = \sum_{n \leqq x} \mu(n) \cdot \frac{x}{n} \log \frac{x}{n} + A_0 \cdot \sum_{n \leqq x} \mu(n) \cdot \frac{x}{n} + O(x) .$$

(17) follows now by (5).

If $f(x)$ is defined for $x > 0$, then

$$\sum_{n \leqq x} \left\{ \frac{x}{n} \cdot \log \frac{x}{n} \cdot f\left(\frac{x}{n}\right) + \frac{x}{n} \cdot \sum_{p^m \leqq x/n} \frac{\log p}{p^m} \cdot f\left(\frac{x}{n \cdot p^m}\right) \right\}$$
$$= \sum_{n \leqq x} \frac{x}{n} \cdot \log \frac{x}{n} \cdot f\left(\frac{x}{n}\right) + \sum_{k \leqq x} \frac{x}{k} \cdot f\left(\frac{x}{k}\right) \cdot \sum_{p^m / k} \log p \qquad (k = n \cdot p^m)$$
$$= \sum_{n \leqq x} \frac{x}{n} \cdot \log \frac{x}{n} \cdot f\left(\frac{x}{n}\right) + \sum_{k \leqq x} \frac{x}{k} \cdot f\left(\frac{x}{k}\right) \cdot \log k = x \cdot \log x \cdot \sum_{n \leqq x} \frac{1}{n} \cdot f\left(\frac{x}{n}\right) .$$

Thus, if we set

$$F(x) \equiv x \cdot \log x \cdot \sum_{n \leqq x} \frac{1}{n} \cdot f\left(\frac{x}{n}\right) ,$$

then, by the inversion formula,

$$x \cdot \log x \cdot f(x) + x \cdot \sum_{p^m \leqq x} \frac{\log p}{p^m} \cdot f(x/p^m) = \sum_{n \leqq x} \mu(n) \cdot F\left(\frac{x}{n}\right) .$$ [1]

In particular, if

$$\sum_{n \leqq x} \frac{1}{n} \cdot f\left(\frac{x}{n}\right) = K + O\left(\frac{\log^k x}{x}\right) ,$$

then

$$\sum_{n \leqq x} \mu(n) \cdot F\left(\frac{x}{n}\right) = K \cdot \sum_{n \leqq x} \mu(n) \cdot \frac{x}{n} \cdot \log\left(\frac{x}{n}\right) + O\left(\sum_{n \leqq x} \log^{k+1}\left(\frac{x}{n}\right)\right) = O(x) ,$$

by (17) and (3), and thus

$$(18) \qquad f(x) \cdot \log x + \sum_{p^m \leqq x} \frac{\log p}{p^m} \cdot f(x/p^m) = O(1) ,$$
$$\textit{if} \quad \sum_{n \leqq x} \frac{1}{n} \cdot f\left(\frac{x}{n}\right) = K + O\left(\frac{\log^k x}{x}\right) .$$

(Selberg's asymptotic formula for $\psi(x)$ corresponds to $f(x) \equiv \psi(x)/x - 1$.) By (9′) and (12′), $f(x) \equiv b_x$ and $f(x) \equiv h_x$ both satisfy the condition of (18), and thus

$$(19) \qquad b_x \cdot \log x + \sum_{p^m \leqq x} \frac{\log p}{p^m} \cdot b_{x/p^m} = O(1)$$

$$(20) \qquad h_x \cdot \log x + \sum_{p^m \leqq x} \frac{\log p}{p^m} \cdot h_{x/p^m} = O(1) .$$

[1] Compare K. Iseki and T. Tatuzawa, "On Selberg's elementary proof of the prime number theorem." Proc. Jap. Acad. 27, 340–342 (1951).

From (16) and (20) it follows that

$$\sum_{n \leqq x} \frac{1}{n} \cdot b_n^2 \geqq |h_x| \cdot \log x + O(1) . \tag{21}$$

If we add to (19)

$$(\log x - A_0) \cdot \log x + \sum_{p^m \leqq x} \frac{\log p}{p^m} \cdot (\log (x/p^m) - A_0) ,$$

which by (1) and (6) is equal to $3/2 \cdot \log^2 x - 3 \cdot A_0 \cdot \log x + O(1)$, we obtain

$$\rho(x) \cdot \log x + \sum_{p^m \leqq x} \frac{\log p}{p^m} \cdot \rho(x/p^m) = \frac{3}{2} \cdot \log^2 x - 3 \cdot A_0 \cdot \log x + O(1) .$$

If $0 < c < 1$, and $c \cdot x < y < x$, then it follows from the last equation that

$$\begin{aligned} \rho(x) \cdot \log x - \rho(y) \cdot \log y &\leqq \frac{3}{2} \cdot (\log^2 x - \log^2 y) + O(1) \\ &= \frac{3}{2} \cdot \log \frac{x}{y} \cdot (\log x + \log y) + O(1) , \end{aligned}$$

$$\log x \cdot (\rho(x) - \rho(y)) + \log \frac{x}{y} \cdot \rho(y) \leqq \frac{3}{2} \cdot \log \frac{x}{y} \cdot (\log x + \log y) + O(1) ,$$

or, since $\rho(y) = \log y + O(1)$,

$$\begin{aligned} \log x \cdot (\rho(x) - \rho(y)) &\leqq \log \frac{x}{y} \cdot \left(\frac{3}{2} \cdot \log x + \frac{1}{2} \cdot \log y\right) + O(1) \\ &< 2 \cdot \log \frac{x}{y} \cdot \log x + O(1) . \end{aligned}$$

Thus

$$\rho(x) - \rho(y) < 2 \cdot \log \frac{x}{y} + O\left(\frac{1}{\log x}\right) ,$$

and, since $\rho(x) = \log x - A_0 + b_x$, it follows that $b_x - b_y < \log x/y + O(1/\log x)$. Also obviously $b_x - b_y \geqq -\log x/y$, because $\rho(x)$ is non-decreasing. Thus we obtain

$$|b_x - b_y| \leqq \log \frac{x}{y} + O\left(\frac{1}{\log x}\right) \quad \text{if } c \cdot x < y < x,\ 0 < c < 1 . \tag{22}$$

## PART III

Let $B \geqq 1$ be an upper bound of $|b_n|$.

Since $b_n - b_{n-1}$ is either $-\log [n/(n-1)]$, or $\log p/n - \log[n/(n-1)]$, it cannot happen that $b_n = b_{n-1} = 0$.

Let the integers $r_1, r_2, \cdots, r_t, \cdots$ be the indices $n$ for which the $b_n$ change signs. Precisely:

$$\text{(23)} \qquad \begin{cases} r_1 = 1\,;\ n = r_t \quad \text{if} \quad b_n \cdot b_{n+1} \leqq 0, \text{ and } b_{n+1} \neq 0\,; \\ \text{if} \quad r_t < v \leqq w < r_{t+1} \text{ then } b_v \cdot b_w > 0\,; \text{ and} \\ |\, b_{r_t} | < (\log r_t)/r_t \text{ for } t > 1. \end{cases}$$

Let $\{s_k\}$ be a sequence of integers, determined as follows: every $r_t$ is an $s_k$; if $\log(r_{t+1}/r_t) < 7 \cdot B$, and $r_t = s_k$, then $r_{t+1} = s_{k+1}$; if $\log(r_{t+1}/r_t) \geqq 7 \cdot B$, enough integers $s_{k+v}$ are inserted between $r_t = s_k$ and $r_{t+1} = s_{k+m}$ such that $3 \cdot B \leqq \log(s_{k+v+1}/s_{k+v}) < 7 \cdot B$, for $v = 0, 1, \cdots, m-1$. If there is a last $r_{t_0} = s_{k_0}$, a sequence $\{s_{k_0+v}\}$ is formed such that $3 \cdot B \leqq \log(s_{k_0+v+1}/s_{k+v}) < 7 \cdot B$. Thus the $s_k$ form a sequence with the following properties:

$$\text{(24)} \qquad \begin{cases} s_1 = 1\,;\ \log(s_{k+1}/s_k) < 7 \cdot B\,;\ \text{for } k > 1, \textit{ either} \\ \log(s_{k+1}/s_k) \geqq 3 \cdot B, \textit{ or } |\, b_{s_k} | \text{ and } |\, b_{s_{k+1}} | \text{ are both} \\ \text{less than } \dfrac{\log s_k}{s_k}\,;\ b_v \cdot b_w > 0 \text{ for } s_k < v \leqq w < s_{k+1}\,. \end{cases}$$

Assume now that $\alpha\,(0 < \alpha < 1/2)$ is such that

$$\text{(25)} \qquad \textit{not} \quad h_x = O(\log^{-\alpha} x).$$

Then $|\, h_x | \cdot \log^\alpha x$ is unbounded. Let $x$ be large, and such that $|\, h_x | \cdot \log^\alpha x \geqq |\, h_y | \cdot \log^\alpha y$ for all $y \leqq x$. Let $c$ and $d$ be positive integers such that

$$\text{(26)} \qquad s_{c-1} < \log x \leqq s_c\,, \quad \text{and} \quad s_d \leqq x < s_{d+1}\,.$$

It will be shown that

$$\frac{1}{2} \cdot (1 - \alpha - o(1)) \cdot S(x) \leqq |\, h_x | \cdot \log x \leqq \frac{1}{3} \cdot (1 + o(1)) \cdot S(x)\,,$$

where

$$\text{(27)} \qquad S(x) \equiv \sum_{k=c+1}^{d} |\, h_{s_k} - h_{s_{k-1}} | \cdot \log(s_k/s_{k-1})\,.$$

From this it will follow that $\alpha \geqq 1/3$.

Clearly

$$|\, h_x | \cdot \log x = |\, h_x | \cdot \log^\alpha x \cdot \left\{ \log^{1-\alpha} x - \log^{1-\alpha} s_d + \sum_{k=2}^{d} (\log^{1-\alpha} s_k - \log^{1-\alpha} s_{k-1}) \right\}$$

$$\geqq \frac{1}{2} \cdot \sum_{k=c+1}^{d} (|\, h_{s_k} | \cdot \log^\alpha s_k + |\, h_{s_{k-1}} | \cdot \log^\alpha s_{k-1}) \cdot (\log^{1-\alpha} s_k - \log^{1-\alpha} s_{k-1})$$

$$\geqq \frac{1}{2} \cdot \sum_{k=c+1}^{d} |\, h_{s_k} - h_{s_{k-1}} | \cdot \log^\alpha s_{k-1} \cdot (\log^{1-\alpha} s_k - \log^{1-\alpha} s_{k-1})\,.$$

If $y < z$, it is easily shown by the mean value theorem that

$$y^\alpha \cdot (z^{1-\alpha} - y^{1-\alpha}) > (1-\alpha) \cdot \frac{y}{z} \cdot (z-y) > \left(1 - \alpha - \frac{z-y}{z}\right)(z-y) .$$

With $y = \log s_{k-1}$, $z = \log s_k$, and from the fact that $s_k > \log x$, $\log (s_k/s_{k-1}) < 7 \cdot B$, it follows by (27) that

$$|h_x| \cdot \log x > \frac{1}{2} \cdot \left(1 - \alpha - \frac{7 \cdot B}{\log \log x}\right) \cdot S(x) . \tag{28}$$

For the next estimate, we need the following lemma.

LEMMA. *Let $v$ and $w$ be positive integers such that*

(1) $\log \dfrac{w}{v} = O(1)$;

(2) $b_n > 0$ *for* $v \leqq n \leqq w$;

(3) $b_v < \dfrac{\log v}{v}$

*Then*

$$\sum_{v \leqq n \leqq w} \frac{1}{n} \cdot b_n^2 \leqq \frac{2}{3} \cdot \log \frac{w}{v} \cdot \sum_{v \leqq n \leqq w} \frac{1}{n} b_n + O\left(\frac{\log(w/v)}{\log v}\right) .$$

*Proof.* If $b_n \leqq 1/3 \cdot \log w/v$ for every $n$ in $[v, w]$, the lemma is obviously correct. Otherwise, let $n_1$ be such that

$$b_{n_1} \geqq \frac{1}{3} \cdot \log \frac{w}{v} , \qquad b_n < \frac{1}{3} \cdot \log \frac{w}{v} \quad \text{for} \quad v \leqq n < n_1 .$$

If $\log (n_1/v) > 1/3 \log (w/v)$, let $z$ $(v \leqq z < n_1)$ be such that $\log (n_1/z) = 1/3 \log (w/v)$; otherwise, let $z = v$. Thus by (22), in every case, $\log (n_1/z) = 1/3 \log (w/v) + O(1/\log v)$. Clearly $b_n - 2/3 \cdot \log w/v < 0$ for $v \leqq n \leqq z$. Thus

$$\begin{aligned} T \equiv & \sum_{v \leqq n \leqq w} \frac{1}{n} \cdot b_n^2 - \frac{2}{3} \cdot \log \frac{w}{v} \cdot \sum_{v \leqq n \leqq w} \frac{1}{n} \cdot b_n \\ \leqq & \sum_{z \leqq n \leqq w} \frac{1}{n} \cdot b_n^2 - \frac{2}{3} \cdot \log \frac{w}{v} \cdot \sum_{z \leqq n \leqq w} \frac{1}{n} \cdot b_n , \end{aligned}$$

$$T \leqq \sum_{z \leqq n \leqq w} \frac{1}{n} \cdot \left(b_n - \frac{1}{3} \cdot \log \frac{w}{v}\right)^2 - \frac{1}{9} \cdot \log^2 (w/v) \cdot \log (w/z) + O\left(\frac{\log (w/v)}{v}\right) .$$

By (22),

$$\begin{aligned} \left| b_n - \frac{1}{3} \log (w/v) \right| &= | b_n - b_{n_1} | + O\left(\frac{\log v}{v}\right) \\ \leqq | \log (n_1/n) | + O\left(\frac{1}{\log v}\right) &= | \log (n_1/z) - \log (n/z) | + O\left(\frac{1}{\log v}\right) , \end{aligned}$$

and thus

$$\left| b_n - \frac{1}{3} \cdot \log \frac{w}{v} \right| \leqq \left| \log \frac{n}{z} - \frac{1}{3} \log \frac{w}{v} \right| + O\left(\frac{1}{\log v}\right) .$$

Thus

$$\begin{aligned} T &\leqq \sum_{z \leqq n \leqq w} \frac{1}{n} \cdot \left( \log \frac{n}{z} - \frac{1}{3} \cdot \log \frac{w}{v} \right)^2 \\ &\quad - \frac{1}{9} \cdot \log^2 (w/v) \cdot \log (w/z) + O\left(\frac{\log (w/v)}{\log v}\right) \\ &= \sum_{z \leqq n \leqq w} \frac{1}{n} \cdot \log^2 (n/z) - \frac{2}{3} \cdot \log \frac{w}{v} \cdot \sum_{z \leqq n \leqq w} \frac{1}{n} \cdot \log \frac{n}{z} + O\left(\frac{\log (w/v)}{\log v}\right) \\ &= \frac{1}{3} \cdot \log^3 (w/z) - \frac{2}{3} \cdot \log^2 (w/v) \cdot \frac{1}{2} \cdot \log^2 (w/z) + O\left(\frac{\log (w/v)}{\log v}\right) , \end{aligned}$$

by (2′), and thus $T \leqq O(\log (w/v)/\log v)$. This completes the proof of the lemma.

COROLLARY 1. *If condition* (3) *is replaced by* $b_w < \log w/w$, *the conclusion still holds; if* $b_n < 0$ *in* $v \leqq n \leqq w$, *the conclusion holds if* $b_n$ *is replaced by* $|b_n|$.

COROLLARY 2. *If instead of* (3) *it is known that* $b_v < \log v/v$ *and* $b_w < \log w/w$ *then*

$$\sum_{v \leqq n \leqq w} \frac{1}{n} \cdot b_n^2 \leqq \frac{1}{3} \cdot \log \frac{w}{v} \cdot \sum_{v \leqq n \leqq w} \frac{1}{n} \cdot |b_n| + O\left(\frac{\log (w/v)}{\log v}\right) .$$

*For a proof, we split* $[v, w]$ *into two intervals by a division point at* $(v \cdot w)^{1/2}$, *and apply the lemma separately to each subinterval.*

COROLLARY 3.

$$\sum_{s_{k-1} < n \leqq s_k} \frac{1}{n} \cdot b_n^2 \leqq \frac{1}{3} \cdot \log (s_k/s_{k-1}) \cdot \sum_{s_{k-1} < n \leqq s_k} \frac{1}{n} \cdot |b_n| + O\left(\frac{\log (s_k/s_{k-1})}{\log s_k}\right) . \tag{29}$$

*Proof.* If $\log (s_k/s_{k-1}) < 3 \cdot B$, this follows from (24) and Corollary 2; if $\log (s_k/s_{k-1}) \geqq 3B$, it is obvious, since $|b_n| \leqq B$.

By (26), $\sum_{n \leqq s_c} 1/n \cdot b_n^2 = O(\log \log x)$, and $\sum_{s_d < n \leqq x} 1/n \cdot b_n^2 = O(1)$; also

$$\sum_{k=c+1}^{d} \frac{\log (s_k/s_{k-1})}{\log s_k} \leqq \sum_{k=c+1}^{d} \log \left(\frac{\log s_k}{\log s_{k-1}}\right) \leqq \log \log x .$$

It follows from (29) that

$$\sum_{n \leqq x} \frac{1}{n} \cdot b_n^2 \leqq \frac{1}{3} \cdot \sum_{k=c+1}^{d} \log (s_k/s_{k-1}) \cdot \sum_{s_{k-1} < n \leqq s_k} \frac{1}{n} \cdot |b_n| + O(\log \log x) .$$

By (8′) $\sum_{s_{k-1} < n \leqq s_k} 1/n \cdot |b_n| = |h_{s_k} - h_{s_{k-1}}| + O(\log s_k/s_k)$, and thus, by (21) and (27),

$$|h_x| \cdot \log x \leqq \frac{1}{3} \cdot S(x) + O(\log\log x) . \tag{30}$$

It follows from (28) and (30) that

$$\left[\frac{1}{3} - \frac{1}{2} \cdot \left(1 - \alpha - \frac{7 \cdot B}{\log\log x}\right)\right] \cdot S(x) \geqq O(\log\log x) ,$$

and since by (25) and (30) $S(x) \geqq K \cdot \log^{1/2} x$, this implies that $\alpha \geqq 1/3$. Thus $h_x = o(\log^{-1/3+\varepsilon} x)$, for every $\varepsilon > 0$, and therefore, by (8′),

$$\sum_{s_{k-1} < n \leqq s_k} \frac{1}{n} \cdot |b_n| = o(\log^{-1/3+\varepsilon} s_k) . \tag{31}$$

In order to find a bound for $|b_x|$, we consider now a particular interval $I_k = (s_{k-1}, s_k]$ ; let us assume that $b_n > 0$ in $I_k$. Let $n_2 \in I_k$ be such that $b_{n_2} \geqq b_n$ for every $n \in I_k$. Let $n_1$ $(s_{k-1} \leqq n_1 < n_2)$ be such that

$$b_{n_1} \leqq \frac{1}{2} \cdot b_{n_2} < b_{n_1+1} .$$

Then

$$\sum_{n \in I_k} \frac{1}{n} \cdot b_n > \sum_{n=n_1+1}^{n_2} \frac{1}{n} \cdot b_n > \frac{1}{2} \cdot b_{n_2} \cdot \log(n_2/n_1) - O(1/s_k) .$$

But by (22),

$$\log(n_2/n_1) \geqq b_{n_2} - b_{n_1} - O\left(\frac{1}{\log s_k}\right) \geqq \frac{1}{2} \cdot b_{n_2} - O\left(\frac{1}{\log s_k}\right) .$$

Thus

$$\sum_{n \in I_k} \frac{1}{n} \cdot b_n > \frac{1}{4} \cdot b_{n_2}^2 - O\left(\frac{1}{\log s_k}\right) .$$

It follows from (31) that $b_{n_2}^2 = o(\log^{-1/3+\varepsilon} n_2)$, and thus

$$b_x = o(\log^{-1/6+\varepsilon} x) . \tag{32}$$

Finally,

$$\begin{aligned}
\psi(x) &= \sum_{n \leqq x} n \cdot (\rho(n) - \rho(n-1)) = [x] \cdot \rho([x]) - \sum_{n \leqq x-1} \rho(n) \\
&= x \cdot (\log x - A_0 + b_x) - \sum_{n \leqq x} (\log n - A_0 + b_n) + O(\log x) \\
&= x \cdot \log x - A_0 \cdot x + b_x \cdot x - x \cdot \log x + x + A_0 \cdot x - \sum_{n \leqq x} b_n + O(\log x) \\
&= x + o(x \cdot \log^{-1/6+\varepsilon} x) + o\left(\sum_{n \leqq x} \log^{-1/6+\varepsilon} n\right) , \quad \text{by (32).}
\end{aligned}$$

The last sum is easily seen to be $o(x \cdot \log^{-1/6+\varepsilon} x)$, and thus

$$\psi(x) = x + o(x \cdot \log^{-1/6+\varepsilon} x) . \tag{33}$$