SIEVE-GENERATED SEQUENCES WITH TRANSLATED INTERVALS

R. G. BUSCHMAN AND M. C. WUNDERLICH

Consider the following sieve process. Let $A^{(1)}$ be the sequence of integers greater than 1. Let $A^{(n+1)}$ be obtained from $A^{(n)} = \{a_1^{(n)}, a_2^{(n)}, \ldots\}$ by eliminating one element from each of the intervals $I_k(n)$, where

$$I_k(n) = \{a_j^{(n)} | n + (k-1)a_n^{(n)} < j \leqslant n + ka_n^{(n)}\}, \qquad k \geqslant 1.$$

We let $a_n = a_n^{(n)}$ and $A = \{a_n\}$ be the sequence of integers that survive the sieve. M. C. Wunderlich **(8)** has found a necessary and sufficient condition for $a_n \sim n \log n$ and, in a more recent paper, M. Wunderlich and W. E. Briggs **(9)** have studied a subclass of the sequences defined above for which $a_n \sim n \log n$. In that paper it was shown that if $f(n)$ is any term whose order lies between $n(\log \log n)^2$ and $n \log n$, then a sieve-generated sequence can be constructed for which $a_n - n \log n \sim f(n)$. However, it was also shown that all sequences generated by the above sieve contain a term whose order is $n(\log \log n)^2$, which is unfortunate in view of the fact that, for primes, $p_n - n \log n \sim n \log \log n$. It is natural to ask whether or not a modification of the above sieve process could produce a more prime-like sequence.

In the sieve of Eratosthenes the first element that is actually sieved out at the kth sieving is p_k^2. Hence there is an interval beyond p_k containing $\pi(p_k^2) - k$ integers, where there is no sieving at all. This property can be incorporated in the above sieve by translating all the intervals $I_k(n)$ by an amount α_n, where α_n is a function of $a_n = a_n^{(n)}$, i.e.

$$I_k(n) = \{a_j^{(n)} | n + \alpha_n + (k-1)a_n^{(n)} < j \leqslant n + \alpha_n + ka_n^{(n)}\}, \qquad k \geqslant 1.$$

If the "prime number theorem" property were to hold for these sequences, then the most interesting value of α to study would be

$$\alpha_k = (\pi(a_k)^2 - k) \sim \tfrac{1}{2}(a_k)^2/\log a_k.$$

In this paper, the authors first prove that $a_n \sim n \log n$ for a large class of α_n, incidentally including $\alpha_n \sim \frac{1}{2}(a_n)^2/\log a_n$. Although the methods used in this paper are not strong enough to prove the existence of a second term, we show that for this "prime-like" α_n, if a second term exists, it must lie between $n(\log \log n)^{1-\epsilon}$ and $n(\log \log n)^{1+\epsilon}$ for any $\epsilon > 0$.

For the present, α_n will be considered an arbitrary non-negative function, and, as we proceed in the proof of the main theorem, we shall impose additional restrictions on α_n as needed. Let $f_n(x)$ denote the number of elements $a_j^{(n)} \leqslant x$

Received February 15, 1966. This research was supported in part by NSF Grant GP 1823 at the University of Colorado.

which are sieved out of $A^{(n)}$ to produce $A^{(n+1)}$, and let $R_n(x)$ denote the number of elements of $A^{(n)}$ not exceeding x. We have $R_{n+1}(x) = R_n(x) - f_n(x)$ and

$$f_k(a_n + 1) = \begin{cases} [(R_k(a_n + 1) - k - \alpha_k)/a_k] + \epsilon_k & \text{if } k + \alpha_k \leqslant n, \\ 0 & \text{otherwise,} \end{cases}$$

where $\epsilon_k = 0$ or 1. Now let $q = q_n$ be defined as the largest k such that $k + \alpha_k \leqslant n$. If $k \leqslant q$, we have a recurrence formula similar to that first used by Hawkins and Briggs **(6)** and later modified **(1, 2, 8, 9)**,

$$R_{k+1}(a_n + 1) = R_k(a_n + 1) - [(R_k(a_n + 1) - k - \alpha_k)/a_k] - \epsilon_k.$$

The procedure can thus follow similar lines.

By iteration we obtain

$$R_{k+1}(a_n + 1) = \sigma_k R_1(a_n + 1) + E_k(a_n + 1),$$

where we have introduced the notations

$$\sigma_k = \prod_{m=1}^{k} (1 - 1/a_m),$$

$$E_k(a_n + 1) = \sum_{m=1}^{k} \frac{\sigma_k}{\sigma_m} \left(\left\{ \frac{R_m(a_n + 1) - m - \alpha_m}{a_m} \right\} + \frac{m + \alpha_m}{a_m} - \epsilon_m \right).$$

Letting $k = q$ and noting that $R_{q+1}(a_n + 1) = n$ we get

(1) $$n = \sigma_q a_n + E_q(a_n + 1).$$

Since $E_q(a_n + 1) > -q > -n$, from (1) we have

(2) $$\sigma_n a_n < \sigma_q a_n = n - E_q(a_n + 1) < 2n.$$

So, summing the above from 1 to n, we obtain

$$\frac{1}{\sigma_n} - 1 = \sum_{k=1}^{n} \left(\frac{1}{\sigma_k} - \frac{1}{\sigma_{k-1}} \right) = \sum_{k=1}^{n} \frac{1}{\sigma_k a_k} > \sum_{k=1}^{n} \frac{1}{2n} > \frac{1}{2} \log n;$$

hence,

(3) $$\sigma_n < 2/(\log n).$$

An upper bound on $E_q(a_n + 1)$ will now be established. First we note that

$$E_q(a_n + 1) < \sum_{k=1}^{q} 1 + \sum_{k=1}^{q} (k + \alpha_k)/a_k.$$

From this point on only those sieves for which the ratio α_k/k is non-decreasing and greater than $1 + \epsilon_1$, $\epsilon_1 > 0$, will be considered. Since $a_1 = 2$ is the first sieving number, $a_k > (2 - \delta)k$ for any $1 > \delta > 0$ and for k large. Hence, since $q_n \to \infty$, we have for n large:

$$\begin{aligned} E_q(a_n + 1) &< q + (2 - \delta)^{-1} \left(\sum_{k=1}^{q} 1 + \sum_{k=1}^{q} \alpha_k/k \right) \\ &\leqslant q + (2 - \delta)^{-1}(q + q \cdot \alpha_q/q) \\ &\leqslant q + n/(2 - \delta). \end{aligned}$$

Now, since $\alpha_q/q > 1$, one can verify that $q < n/(2 + \epsilon_1)$. Hence, by choosing a sufficiently small δ above ($\delta < \epsilon_1/(1 + \epsilon_1)$), one obtains

$$E_q(a_n + 1) < (1 - \epsilon)n \tag{4}$$

for some $0 < \epsilon < 1$ and for n large. If we now assume that $\log \alpha_n/\log n < r$ for some positive constant r, it can then be shown that

$$\log n/\log q < r(1 + o(1))$$

for large n. The inequalities (3) and (4) produce

$$\begin{aligned} n = R_q(a_n + 1) &= \sigma_q a_n + E_q(a_n + 1) \\ &< 2a_n/(\log q) + (1 - \epsilon)n \\ &< 2r(1 + o(1))a_n/(\log n) + (1 - \epsilon)n. \end{aligned}$$

Hence, for a suitable positive constant c_1,

$$a_n > c_1 n \log n. \tag{5}$$

To get an upper bound for a_n we estimate σ_n/σ_q as follows:

$$\begin{aligned} \frac{\sigma_n}{\sigma_q} &= \prod_{k=q+1}^{n}\left(1 - \frac{1}{a_k}\right) > \prod_{k=q+1}^{n}\left(1 - \frac{1}{c_1\, k \log k}\right) \\ &\geqslant \exp\left(\sum_{k=q+1}^{n} \frac{-1 + o(1)}{c_1\, k \log k}\right) \\ &> \exp\left(\frac{-1 + o(1)}{c_1} \log \frac{\log n}{\log q}\right) \\ &> \exp\left(\frac{-1 + o(1)}{c_1} \log r\right) \\ &\geqslant r^{(-1+o(1))/c_1} = c_2 > 0. \end{aligned} \tag{6}$$

Hence, using (4), we get

$$\frac{1}{\sigma_n\, a_n} = \frac{1}{\sigma_n\, a_n}\, \frac{1}{\sigma_n/\sigma_q} < \frac{1}{c_2}\left(\frac{1}{n - E_q(a_n + 1)}\right) < \frac{1}{\epsilon c_2 n} = \frac{c_3}{n}.$$

We now sum the above from 1 to n in a telescopic series to obtain

$$1/\sigma_n - 1 < c_3 \log n$$

and then use (2) to obtain

$$a_n < c_4\, n \log n$$

for some constant $c_4 > 0$. We have proved the following theorem.

THEOREM 1. *If a_n is a sieve-generated sequence for which $\alpha_n/n > 1 + \epsilon$, $\epsilon > 0$, α_n/n is non-decreasing, and $\log \alpha_n/\log n < r$, where $r > 0$, then there exist two positive constants c_1 and c_4 such that*

$$c_1 < a_n/(n \log n) < c_4.$$

We shall now proceed to estimate the constants c_1 and c_4. To do this we restrict our attention to the case where

$$\alpha_n \sim c(a_n)^a(\log a_n)^b,$$

where a, b, and c are constants such that $a > 1$ and $c > 0$. Although α_n is a function of n, it is desirable to regard it as a function of a_n, since in the sieve of Eratosthenes the interval where no sieving takes place has length

$$\pi(p_n^2) - n \sim \tfrac{1}{2}p_n^2(\log p_n)^{-1}.$$

One first uses the result that $a_n > c_1 n \log n$ and, in the same manner that (4) was obtained, now produces the improved estimate

$$E_q(a_n + 1) = o(n). \tag{7}$$

Next (2) becomes $(\sigma_n a_n)^{-1} > (1 + o(1))$ and hence (3) becomes

$$\sigma_n < (1 + o(1)) \log n.$$

Since a_n is of the same order of magnitude as $n \log n$, and α_n is of the same order of magnitude as $n^a(\log n)^{a+b}$, one can then show that $q = q_n$ is of the same order of magnitude as $n^{1/a}(\log n)^{-(a+b)/a}$. Hence, $\log q \sim a^{-1} \log n$ and

$$\sigma_q < (1 + o(1))/\log q = (1 + o(1))a/\log n.$$

We can now improve (5) as follows:

$$\begin{aligned} n \leqslant R_q(a_n + 1) &= \sigma_q a_n + o(n) \\ &< (1 + o(1))a a_n/\log n + o(n), \end{aligned}$$

which gives us for n large

$$a_n > a^{-1}(1 + o(1))n \log n. \tag{8}$$

Now using (6) and (8) we obtain, successively,

$$\begin{aligned} \sigma_n/\sigma_q &> a^{-a+o(1)} = (1 + o(1))a^{-a}, \\ (\sigma_n a_n)^{-1} &< (1 + o(1))a^a/n, \\ \sigma_n^{-1} &< (1 + o(1))a^a \log n, \end{aligned}$$

and, thus,

$$\sigma_q^{-1} < (1 + o(1))a^a \log q = (1 + o(1))a^{a-1} \log n.$$

Then from (1) and (7) we have $\sigma_q a_n = n + o(n)$ so that

$$a_n < (1 + o(1))a^{a-1} n \log n.$$

It is interesting to note that in the "prime-like" case where $a = 2$, $b = -1$, and $c = \frac{1}{2}$ we have, at this point, the "Chebyshev Theorem"

$$\tfrac{1}{2} - \epsilon < a_n/(n \log n) < 2 + \epsilon.$$

The method just described of establishing upper and lower bounds on $a_n/(n \log n)$ is now iterated to obtain sharper bounds. It will be shown ultimately that these bounds can be made arbitrarily close to 1, proving that $a_n \sim n \log n$. However, at each iteration terms which are $o(1)$ are introduced in the bounds which would cause difficulty when passing to the limit. Therefore the proof of the following lemma is presented in detail.

LEMMA 2.1. *Suppose that for every* $\epsilon > 0$, *there exists an* $N = N(\epsilon)$ *such that for every* $n > N$, $a_n < (\tau + \epsilon)n \log n$ *for a given constant* τ. *Then for any* $\delta > 0$, *there exists an* $M = M(\delta)$ *such that for all* $m > M$

$$a_m > (\tau_1 - \delta) m \log m,$$

where

$$\tau_1 = a^{1/\tau - 1}.$$

Proof. Since $q_n \to \infty$ as $n \to \infty$, one can fix M large enough so that for all $n > M$, $a_q < (\tau + \epsilon) q \log q$. Then for $n > M$,

$$\frac{\sigma_n}{\sigma_q} = \prod_{k=q+1}^{n} \left(1 - \frac{1}{a_n}\right) < \prod_{k=q+1}^{n} \left(1 - \frac{1}{(\tau + \epsilon) n \log n}\right)$$
$$\leqslant \exp\left\{\sum_{k=q+1}^{n} \log\left(1 - \frac{1}{(\tau + \epsilon) n \log n}\right)\right\}.$$

(Throughout this proof we shall make use of the symbols $\epsilon(n)$ and $\epsilon_1(n)$ to denote functions of n which tend to zero for n large. These functions will be modified throughout the proof without changing notation.) From the expansion of the logarithm we have

$$\frac{\sigma_n}{\sigma_q} < \exp\left(-\sum_{k=q+1}^{n} \frac{1 + \epsilon(k)}{(\tau + \epsilon) k \log k}\right). \tag{9}$$

The second term is treated separately. Let $\epsilon(n)$ be the largest of the $\epsilon(k)$ in the range (q_n, n), so that

$$\sum_{k=q+1}^{n} \frac{\epsilon(k)}{(\tau + \epsilon) k \log k} < \epsilon(n) \sum_{k=q+1}^{n} \frac{1}{(\tau + \epsilon) k \log k}.$$

Using this in (9) above, we get

$$\frac{\sigma_n}{\sigma_q} < \exp\left\{\left(\frac{-1 + \epsilon(n)}{\tau + \epsilon}\right) \sum_{k=q+1}^{n} \frac{1}{k \log k}\right\}$$
$$= \exp\left\{\left(\frac{-1 + \epsilon(n)}{\tau + \epsilon}\right) \log\left(\frac{\log n}{\log q}\right)\right\}.$$

Note that a constant is introduced in the estimate of the sum but we absorb it into $\epsilon(n)$. By virtue of our definition of α_n, $(\log n)/(\log q) = a + \epsilon_1(n)$, where $\epsilon_1(n) \to 0$, so that

$$\frac{\sigma_n}{\sigma_q} < \exp\left\{\left(\frac{-1+\epsilon(n)}{\tau+\epsilon}\right)\log\,(a+\epsilon_1(n))\right\}$$
$$= \exp\left\{\left(\frac{-1+\epsilon(n)}{\tau+\epsilon}\right)(\log a+\epsilon_1(n))\right\},$$

by redefining $\epsilon_1(n)$. We can absorb $\epsilon_1(n)$ into $\epsilon(n)$ by redefining $\epsilon(n)$ and we get

$$\frac{\sigma_n}{\sigma_q} < \exp\left\{\left(\frac{-1+\epsilon(n)}{\tau+\epsilon}\right)\log a\right\}$$
$$= a^{-1/(\tau+\epsilon)}a^{\epsilon(n)/(\tau+\epsilon)}.$$

The second factor, which tends to 1, can be replaced by $1+\epsilon(n)$ for $\epsilon(n)$ possibly redefined. Hence

$$\sigma_n/\sigma_q < a^{-1/(\tau+\epsilon)}(1+\epsilon(n)) \qquad \text{for } n > N(\epsilon). \tag{10}$$

Since ϵ is an arbitrary positive constant, we may replace ϵ in (10) by $\epsilon_1(n)$ to make (10) valid for all n. Then

$$\sigma_n/\sigma_q < a^{-1/\tau}a^{\epsilon_1(n)/\tau(\tau+\epsilon_1(n))}(1+\epsilon(n)).$$

But for a new $\epsilon_1(n)$

$$a^{\epsilon_1(n)/\tau(\tau+\epsilon_1(n))}(1+\epsilon_1(n));$$

hence

$$\sigma_n/\sigma_q < a^{-1/\tau}(1+\epsilon_2(n)), \tag{11}$$

where $\epsilon_2(n) = (1+\epsilon(n))(1+\epsilon_1(n)) - 1$.

Now using (11) we obtain

$$\sigma_n a_n = \sigma_q a_n(\sigma_n/\sigma_q) < \sigma_q a_n a^{-1/\tau}(1+\epsilon(n)).$$

Since $\sigma_q\,a_n = n - E_q(a_n+1) = n(1+\epsilon_1(n))$, we have, after combining $\epsilon_1(n)$ and $\epsilon(n)$,

$$(\sigma_n\,a_n)^{-1} > a^{1/\tau}(1+\epsilon(n))/n, \tag{12}$$

which is true for all n. We now sum (12) from 1 to n to obtain

$$\sum_{k=1}^{n}(\sigma_k\,a_k)^{-1} > \sum_{k=1}^{n} a^{1/\tau}(1+\epsilon(k))/k$$
$$= a^{1/\tau}(\log n + \epsilon(n)\log n)$$

for suitably redefined $\epsilon(n)$. However, since this sum telescopes, we get

$$1/\sigma_n > (1+\epsilon(n))(a^{1/\tau}\log n)$$

or

$$1/\sigma_q > (1+\epsilon(q))(a^{1/\tau}\log q),$$

and, since $q \to \infty$ as $n \to \infty$, $\epsilon(q)$ can be replaced by an $\epsilon(n)$. Furthermore by our choice of α_n,

$$\log q = a^{-1} \log n(1 + \epsilon_1(n))$$

and so, for a redefined $\epsilon(n)$, we have

$$1/\sigma_q > (1 + \epsilon(n))a^{1/\tau-1}(\log n). \tag{13}$$

Now, since $n = \sigma_q a_n + E_q(a_n + 1) = \sigma_q a_n - n\epsilon_1(n)$, we have

$$\sigma_q a_n = n(1 + \epsilon_1(n)).$$

Hence, multiplying this with (13) and for a new $\epsilon(n)$, we get

$$a_n > a^{1/\tau-1}(n \log n)(1 + \epsilon(n)).$$

Since $a > 0$, this can be rewritten by redefining $\epsilon(n)$ by

$$a_n > (a^{1/\tau-1} + \epsilon(n))n \log n.$$

Now, to complete the proof, let $\delta > 0$ be given. There exists an M such that for all $n > M$, $\epsilon(n) - \delta$; hence,

$$a_n > (\tau_1 - \delta)n \log n.$$

LEMMA 2.2. *Suppose that for every $\epsilon > 0$ there exists an $N = N(\epsilon)$ such that for every $n > N$, $a_n > (\tau_1 - \epsilon)n \log n$ for a given constant τ_1. Then for any $\delta > 0$, there exists an $M = M(\delta)$ such that for $m > M$,*

$$a_m < (\tau_2 + \delta)m \log m,$$

where

$$\tau_2 = a^{1/\tau_1 - 1}.$$

The proof of this lemma is exactly parallel to the proof of Lemma 2.1.

We now define a sequence of real numbers $\{x_n\}$ as follows. For $a > 1$ we let $x_1 = a^{-1}$ and we let $x_{n+1} = a^{1/x_n - 1}$, $n = 2, 3, \ldots,$

LEMMA 2.3. *For* $0 < a \leqslant e$, $\operatorname{lin}_{n\to\infty} x_n = 1$.

Proof. Let $y_n = (ax_n)^{-1}$, $a^a = e^\beta$, so that the iteration takes the form $y_{n+1} = \exp(-\beta y_n)$. It can be shown that $y_1 = 1$, $y_{2n+1} < y_{2n-1}$; $y_2 = e^{-\beta} < 1$, $y_{2n} > y_{2n-2}$; and $y_{2n+1} > y_{2m}$ for all $m, n \geqslant 1$. Since the limit must satisfy the equation $L = \exp(-\beta e^{-\beta L})$, which has a unique solution for $0 < \beta \leqslant e$, we find that $y_n \to a^{-1}$; hence $x_n \to 1$.

LEMMA 2.4. *Let $\epsilon > 0$. If n is odd, there exists an $N = N(\epsilon, n)$ such that for all $k > N$,*

$$a_k > (x_n - \epsilon)k \log k.$$

If n is even,

$$a_k < (x_n + \epsilon)k \log k.$$

https://doi.org/10.4153/CJM-1967-049-4 Published online by Cambridge University Press

Proof. This is by induction on n. It follows from (8) that the lemma is true for $n = 1$. Suppose it is true for n, where n is odd. This satisfies the hypothesis of Lemma 2.2 where $\tau_1 = x_n$. Then the conclusion of Lemma 2.2 proves Lemma 2.4 for $n + 1$ since $\tau_2 = a^{1/\tau_1 - 1} = x_{n+1}$. For even n the conclusion follows from Lemma 2.1.

From the above lemmas we now have the following theorem.

THEOREM 2. *If* $\{a_n\}$ *is a sieve-generated sequence where*

$$\alpha_n = c(a_n)^a(\log a_n)^b - n$$

with a, b, *and* c *constants such that* $1 < a < e$ *and* $c > 0$, *then* $a_n \sim n \log n$.

Proof. Suppose that $a \nsim n \log n$. Then there exists an $\epsilon > 0$ such that either
(a) for infinitely many n, $a_n > (1 + \epsilon)n \log n$, or
(b) for infinitely many n, $a_n < (1 - \epsilon)n \log n$.
We shall show that (a) cannot hold, and that a similar proof works for (b). Since $x_n \to 1$, there exists $N(\epsilon)$ such that for all $n > N$, $x_n < 1 + \epsilon/2$. By Lemma 2.4 there exists an $M(\epsilon, N)$ such that, for all $k > M$ and for some $n > N$,

$$a_k < (x_n + \epsilon/2)k \log k < (1 + \epsilon)k \log k,$$

which contradicts (a).

In another paper **(2)** we have considered sieving intervals of length μ_n (a function of $a_n^{(n)}$); that is, one element is eliminated from each of the intervals

$$I_k(n) = \{a_j^{(n)} | n + (k - 1)\mu_n < j \leqslant n + k\mu_n\}, \qquad k \geqslant 1.$$

For these untranslated intervals and $\mu_n = \lambda a_n^{(n)}$, for some constant λ, we obtained $a_n \sim \lambda^{-1} n \log n$. The case of translated intervals can also be handled in a manner similar to that developed in this paper.

We shall now turn our attention to the investigation of the second term. Throughout the discussion we continue with the assumption that

$$\alpha_n = (c + o(1))(a_n)^a(\log a_n)^b.$$

Using this in the relationship

$$q + \alpha_q \leqslant n \leqslant q + 1 + \alpha_{q+1},$$

we obtain

$$n = (c + o(1))q^a(\log q)^{a+b}, \tag{14}$$

and solving for q yields

$$q = \{(1 + o(1))a^{a+b}c^{-1}n(\log n)^{-a-b}\}^{1/a}. \tag{15}$$

The iterative procedure produces the result

$$\sigma_q = (1 + o(1))/(\log n)$$

and, hence,

$$\sigma_n = (1 + o(1))/(a \log n).$$

From this we prove the following lemma.

LEMMA 3.1. $E_q(a_n + 1) = (1 + o(1))n/(\log n)$.

Proof. If the values for a_n, α_n, and σ_q are substituted into the definition of $E_q(a_n + 1)$, then we have

$$\begin{aligned} E_q(a_n + 1) &= \frac{1 + o(1)}{\log n}\left(\sum_{k=k_0}^{q} (1 + o(1))\log k \right. \\ &\quad \left. + \sum_{k=k_0}^{q} (c + o(1))ak^{a-1}(\log k)^{a+b}\right) \\ &= \frac{1 + o(1)}{\log n}\left(o(n) + ca\sum_{k=k_0}^{q} k^{a-1}(\log k)^{a+b}\right) \\ &= \frac{1 + o(1)}{\log n}\,[o(n) + ca\{a^{-1}(1 + o(1))q^a(\log q)^{a+b}\}] \\ &= (1 + o(1))n/(\log n) \end{aligned}$$

from (14).

In order to obtain a second term for a_n we need a second term for σ_n/σ_q and this, in turn, relies on a second term for a_n. Therefore we cannot obtain a second term using these methods. However, this relationship between the second terms of a_n and σ_n/σ_q suggests that if certain assumptions are made concerning the existence of a second term, we can obtain results concerning its value.

LEMMA 3.2. *If*

$$a_n - n\log n = A(1 + o(1))n(\log\log n)^B,$$

where A and B are constants, then

$$\sigma_n/\sigma_q = a^{-1}[1 + (1 + o(1))\{A(a - 1)(\log\log n)^B - (a + b)\log\log n\}/(\log n)].$$

Proof.

$$\begin{aligned} \sigma_n/\sigma_q &= \exp\left(\sum_{k=q+1}^{n} \log(1 - 1/a_k)\right) \\ &= \exp\left\{\sum_{k=q+1}^{n} \log\left(1 - \frac{1}{k\log k} + (1 + o(1))\frac{A(\log\log k)^B}{k\log^2 k}\right)\right\} \\ &= \exp\left(-\sum_{k=q+1}^{n} \frac{1}{k\log k} + (1 + o(1))A\sum_{k=q+1}^{n} \frac{(\log\log k)^B}{k\log^2 k}\right). \end{aligned}$$

The second sum can be estimated as follows, using (15):

$$\begin{aligned} \sum_{k=q+1}^{n} \frac{(\log\log k)^B}{k\log^2 k} &= \int_q^n \frac{(\log\log x)^B}{x\log^2 x}\,dx + O\left(\frac{(\log\log q)^B}{q\log^2 q}\right) \\ &= \left(\frac{(\log\log q)^B}{\log q} - \frac{(\log\log n)^B}{\log n}\right)\left(1 + O\left(\frac{1}{\log\log n}\right)\right) \\ &= (1 + o(1))(a - 1)(\log\log n)^B/(\log n). \end{aligned}$$

https://doi.org/10.4153/CJM-1967-049-4 Published online by Cambridge University Press

The first sum is readily evaluated so that if we use the expression for $(\log q)/(\log n)$ and estimate the remainder of the exponential with

$$\exp u = 1 + u(1 + o(1)),$$

we have

$$\begin{aligned}\frac{\sigma_n}{\sigma_q} &= \exp\left\{\log\left(\frac{\log q}{\log n}\right) + (1 + o(1))A(a-1)\frac{(\log\log n)^B}{\log n}\right\}\\ &= a^{-1}\{1 - (a+b)(1+o(1))(\log\log n)/(\log n)\}\\ &\qquad\times\{1 + A(a-1)(1+o(1))(\log\log n)^B/(\log n)\}\\ &= a^{-1}[1 + (1+o(1))\{A(a-1)(\log\log n)^B - (a+b)\log\log n\}/(\log n)].\end{aligned}$$

One can now use Lemma 3.2 to compute $(a_n\sigma_n)^{-1}$, where

$$a_n\sigma_q = n - (1 + o(1))n/(\log n)$$

from Lemma 3.1. Then σ_n^{-1}, σ_q^{-1}, and a_n can be successively computed and the result compared with the original assumption about a_n. The pattern of computations is the same as in the proof of Theorem 2 so that only the following expressions will be given:

$$\begin{aligned}\frac{1}{a_n\sigma_n} &= \frac{a}{n}[1 - (1+o(1))\{A(a-1)(\log\log n)^B\\ &\qquad - (a+b)(\log\log n) - 1\}/(\log n)],\\ \frac{1}{\sigma_n} &= a\log n - a(1+o(1))\left(\frac{A(a-1)}{B+1}(\log\log n)^{B+1}\right.\\ &\qquad\left. - \frac{a+b}{2}(\log\log n)^2 - \log\log n\right),\\ \frac{1}{\sigma_q} &= \log n - a(1+o(1))\left(\frac{A(a-1)}{B+1}(\log\log n)^{B+1}\right.\\ &\qquad\left. - \frac{a+b}{2}(\log\log n)^2 + \frac{b}{a}\log\log n\right),\\ a_n &= n\log n - an(1+o(1))\left(\frac{A(a-1)}{B+1}(\log\log n)^{B+1}\right.\\ &\qquad\left. - \frac{a+b}{2}(\log\log n)^2 + \frac{b}{a}\log\log n\right).\end{aligned} \tag{16}$$

We shall now examine three cases.

Case 1. $B = 1$. The resulting form becomes

$$a_n = n\log n - \tfrac{1}{2}an(1+o(1))(A(a-1) - (a+b))(\log\log n)^2,$$

where the second term is of too high an order of magnitude unless

$$A = (a+b)/(a-1).$$

https://doi.org/10.4153/CJM-1967-049-4 Published online by Cambridge University Press

The case $a = 2$, $b = -1$ yields $A = 1$ so that the prime-like case is not ruled out.

Case 2. $B > 1$. The form for a_n which results is

$$a_n = n \log n - (1 + o(1))(A(a-1)/(B+1))(\log \log n)^{B+1},$$

where the second term is now always of too high an order of magnitude.

Case 3. $B < 1$. The resulting form for a_n is

$$a_n = n \log n + \tfrac{1}{2}a(a+b)(1+o(1))n(\log \log n)^2$$

which is also too large if $a + b \neq 0$. If $a + b = 0$, we have a contradiction unless $B = 0$ in which case $A = (a-1)^{-1}$ must hold.

Further possible forms for the second term of a_n can be eliminated by repeating the above arguments with inequalities.

LEMMA 3.3. *If* $B > 1$ *and* $a_n - n \log n > (\log \log n)^B$ *for* n *sufficiently large, then*

$$\sigma_n/\sigma_q > a^{-1}[1 + (1+o(1))\{A(a-1)(\log \log n)^B - (a+b)\log \log n\}/(\log n)].$$

Proof. Repeat the proof of Lemma 3.2, substituting $>$ for $=$ throughout.

We can now repeat the arguments in (16) substituting the appropriate inequality for the equality and obtain

$$a_n - n \log n < -A(a-1)a(B+1)^{-1}(1+o(1))n(\log \log n)^{B+1}.$$

Comparing this with the hypothesis of Lemma 3.3 we arrive at a contradiction; thus, we may eliminate all second terms whose order of magnitude is $\geqslant n(\log \log n)^{1+\epsilon}$. In a completely analogous way, one can eliminate terms whose order of magnitude is $\leqslant n(\log \log n)^{1-\epsilon}$, except for the exceptional cases already noted concerning $B = 0$.

Although the results here cited are negative and fragmentary, there is ample evidence to support the conjecture that if a second term exists, it must be of the form

$$\frac{a+b}{a-1} n \log \log n,$$

which is the prime-like case for $a = 2$, $b = -1$.

Added in proof: The authors would like to thank Martin Bates for correcting an error in the proof of Lemma 2.3.

REFERENCES

1. W. E. Briggs, *Prime-like sequences generated by a sieve process*, Duke Math. J., *30* (1963), 297–312.

2. R. G. Buschman and M. C. Wunderlich, *Sieves with generalized intervals*, Boll. Un. Mat. Ital. (to appear).

3. P. Erdos and E. Jabotinsky, *On sequences generated by a sieving process*, Nederl. Akad. Wetensch. Indag. Math., *61* (1958), 115–128.

4. V. Gardiner, R. Lazarus, M. Metropolis, and S. Ulam, *On certain sequences of integers defined by sieves*, Math. Mag., *29* (1955/56), 117–122.

5. D. Hawkins, *The random sieve*, Math. Mag., *31* (1957/58), 1–3.

6. D. Hawkins and W. E. Briggs, *The lucky number theorem*, Math. Mag., *31* (1957/58), 277–280.

7. B. Lachapelle, *L'Esperance mathématique du nombre de nombres premiers aléatoires inférieurs ou egual à x*, Can. Math. Bull., *4* (1961), 139–142.

8. M. C. Wunderlich, *Sieve-generated sequences*, Can. J. Math., *18* (1966), 291–299.

9. M. C. Wunderlich and W. E. Briggs, *Second and third term approximations of sieve-generated sequences*, Illinois J. Math., *10* (1966), 694–700.

State University of New York,
Buffalo