The Prime Number Graph

By Carl Pomerance

Abstract. Let $p_n$ denote the $n$th prime. The prime number graph is the set of lattice points $(n, p_n)$, $n = 1, 2, \ldots$. We show that for every $k$ there are $k$ such points that are collinear. By considering the convex hull of the prime number graph, we show that there are infinitely many $n$ such that $2p_n < p_{n-i} + p_{n+i}$ for all positive $i < n$. By a similar argument, we show that there are infinitely many $n$ for which $p_n^2 > p_{n-p_{n+i}}$ for all positive $i < n$, thus verifying a conjecture of Selfridge. We make some new conjectures.

1. Introduction. Let $p_n$ denote the $n$th prime. In [3], Erdös reports that he and E. G. Straus conjecture that for all sufficiently large $n$ there is a positive $i < n$ with $p_n^2 < p_{n-p_{n+i}}$. In [4], Erdös repeats this conjecture, but here it is remarked that J. L. Selfridge takes the opposite opinion. That is, Selfridge conjectures that there are infinitely many $n$ with

$$p_n^2 > p_{n-p_{n+i}}$$

for all positive $i < n$.

In this paper we give a surprisingly short proof of Selfridge's conjecture using only that $\log p_n = O(n)$. Our proof is geometric and generalizes to prove a number of related results. In particular, by studying the geometry of the prime number graph; that is, the set of points $(n, p_n)$ for $n = 1, 2, \ldots$, we show there are infinitely many $n$ for which

$$2p_n < p_{n-i} + p_{n+i}$$

for all positive $i < n$.

Although introduced as a tool to prove (1.2), the prime number graph is also related to other problems. For example, the famous $\pi(x + y)$ conjecture of Hardy and Littlewood ($\pi(x + y) \leq \pi(x) + \pi(y)$ for $2 \leq x, y$) can be recast into a statement about the prime number graph (cf. Segal [15]). Of course, the $\pi(x + y)$ conjecture is now in disrepute (Hensley and Richards [8]), but it has not yet been disproved.

We conjecture that for every $k$ there are $k$ points on the prime number graph which form an arithmetic progression of vectors. This conjecture will follow if for every $k$ there is an $n$ with $p_{n+1}, p_{n+2}, \ldots, p_{n+k}$ all in arithmetic progression. This last statement can be shown to follow from the prime $k$-tuples hypothesis (another conjecture of Hardy and Littlewood, this one not in disrepute—see [8] for a statement). In this paper we give an unconditional proof of the following weaker result:
THEOREM. For every $k$ there are $k$ distinct integers $n(1), \ldots, n(k)$ such that $(n(1), p_{n(1)}), \ldots, (n(k), p_{n(k)})$ are collinear.

An old and well-known conjecture is that for every $k$ there are $k$ primes in arithmetic progression. This conjecture, although similar sounding to the above theorem, neither implies it nor is implied by it.

I have benefitted greatly in preparing this paper from stimulating conversations with many of my colleagues. In particular, I gratefully acknowledge the help of E. Azoff, E. R. Canfield, H. Pittie, and J. Shyr. A quick look at the references is all that is needed to see the enormous influence of P. Erdős on the topics dealt with in this paper. I would like to especially thank him for the encouragement he has given me.

2. The Convex Hull. In this section we prove that (1.1) and (1.2) each hold for infinitely many $n$ and prove some related results.

THEOREM 2.1. Let $0 < a_1 < a_2 < \ldots$ be a sequence of numbers with $\lim n/a_n = 0$. Then there are infinitely many $n$ for which

\[ 2a_n < a_{n-1} + a_{n+1} \text{ for all positive } i < n. \]

Proof. The boundary of the convex hull of the set $\{(n, a_n): n = 1, 2, \ldots\}$ is polygonal. The assumption $\lim n/a_n = 0$ implies that the nonvertical portion of this polygonal boundary is convex and has infinitely many vertices. Moreover, each of these vertices is in the form $(n, a_n)$ for some $n$. It immediately follows that (2.1) holds for each such $n$. □

COROLLARY. There are infinitely many $n$ for which (1.2) holds.

THEOREM 2.2. Let $0 < a_1 < a_2 < \ldots$ be a sequence of numbers with $\lim a_n/n = 0$. Then there are infinitely many $n$ for which

\[ 2a_n > a_{n-1} + a_{n+1} \text{ for all positive } i < n. \]

Proof. The proof is the same as that for Theorem 2.1, except that now the non-horizontal portion of the convex hull boundary is concave and has infinitely many vertices. □

COROLLARY. There are infinitely many $n$ for which (1.1) holds.

Proof. We apply Theorem 2.2 to the sequence $\log p_1, \log p_2, \ldots$. By either the prime number theorem or the more elementary estimates of Chebyshev, we have $p_n < cn \log n$ where $c$ is a constant. Then $\lim (\log p_n)/n = 0$. We, thus, have infinitely many $n$ for which

\[ 2 \log p_n > \log p_{n-1} + \log p_{n+1} \text{ for all positive } i < n. \]

Exponentiating, we have (1.1). □

If $0 < \alpha < 1$, we can similarly consider the sequence $p_1^\alpha, p_2^\alpha, \ldots$, and prove that the set $P(\alpha)$ of $n$ for which

\[ 2p_n^\alpha > p_{n-1}^\alpha + p_{n+1}^\alpha \text{ for all positive } i < n, \]

is infinite. If $P(0)$ is the set of $n$ for which (1.1) holds, we have

\[ 0 \leq \alpha \leq \beta < 1 \text{ implies } P(\beta) \subset P(\alpha). \]
To see this, let \( f(x) = r^x + s^x \), where \( r = p_{n-i}/p_n \), \( s = p_{n+i}/p_n \). Then \( f \) is convex, \( f(0) = 2, f'(0) = \log(rs) \). Hence, if \( f(\beta) < 2 \) for some \( \beta > 0 \), then \( rs < 1 \) and \( f(\alpha) < 2 \) for all \( \alpha, 0 < \alpha \leq \beta \).

As usual, let \( \li(x) \) denote the improper integral

\[
\int_0^x \frac{dt}{\log t}.
\]

**Theorem 2.3.** There are infinitely many \( n \) such that

(2.5) \( 2 \li(p_n) < \li(p_{n-i}) + \li(p_{n+i}) \) for all positive \( i < n \);

and there are infinitely many \( n \) such that

(2.6) \( 2 \li(p_n) > \li(p_{n-i}) + \li(p_{n+i}) \) for all positive \( i < n \).

**Proof.** It is known (the prime number theorem) that \( \li(p_n) \sim n \) as \( n \to \infty \). From a famous result of Littlewood (cf. Ingham [9]) we have

\[
\lim \sup (\li(p_n) - n) = +\infty, \quad \lim \inf (\li(p_n) - n) = -\infty.
\]

(The theorem to which we refer is that \( \li(x) - \pi(x) \) is unbounded from above and below. It is an easy corollary that the same is true when \( x \) is restricted to prime values.) Hence, there are infinitely many vertices on the upper boundary of the convex hull of \( \{(n, \li(p_n)): n = 1, 2, \ldots \} \) and also infinitely many vertices on the lower boundary. The upper boundary vertices give solutions of (2.6) while the lower boundary vertices give solutions of (2.5). \( \square \)

Given Littlewood’s result quoted above on the gross swings in magnitude of \( \li(x) - \pi(x) \), it would seem that if \( f(x) \) is any concave function on \( x > 0 \), then \( |\pi(x) - f(x)| \) is unbounded (or equivalently, if \( g(x) \) is any convex function on \( x > 0 \), then \( |p_n - g(n)|/\log 2n \) is unbounded). However, this conclusion does not follow from Littlewood’s theorem; that is, there are hypothetical prime distributions which are consistent with Littlewood’s theorem and for which \( |\pi(x) - f(x)| \) is bounded for some concave function \( f(x) \). An analogy might be in order. The orbit of the moon around the sun certainly has gross discrepancies from the orbit of the earth around the sun; but nevertheless, the moon’s orbit is always concave with respect to the sun. However, we can prove

**Theorem 2.4.** If \( f(x) \) is any concave function on \( x > 0 \), then \( |\pi(x) - f(x)| \) is unbounded. Also, if \( g(x) \) is any convex function on \( x > 0 \), then \( |p_n - g(n)|/\log 2n \) is unbounded.

**Proof.** By the prime number theorem and a simple computation, the second statement is equivalent to the first, so we need only prove the first. Let \( f(x) \) be a concave function on \( x > 0 \) such that \( |\pi(x) - f(x)| \) is bounded. We may assume \( f \) is smooth. Since \( f(x) \sim x/\log x \), an easy argument gives \( f'(x) \sim 1/\log x \). (By the mean value theorem and \( f' \) nonincreasing we have

\[
\frac{f(2x) - f(x)}{x} \leq f'(x) \leq \frac{f(x) - f(x/2)}{x/2}.
\]
for all \( x > 0 \). Multiplying this inequality by \( \log x \) and letting \( x \to \infty \), we get 
\[
\lim_{x \to \infty} (\log x) f'(x) = 1.
\]
Thus, for any constant \( c \), 
\[
\lim_{x \to \infty} (f(x + c \log x) - f(x)) = c.
\]

Now an old result of Westzynthius (see [1]) is that 
\[
\lim \sup (p_{n+1} - p_n)/\log n = \infty.
\]
We, thus, have for every \( c \), arbitrarily large numbers \( x_0 \) for which \( \pi(x) \) is constant on the interval \([x_0, x_0 + c \log x_0]\). For large \( x_0 \), there is then in this interval an \( x \) where \( |\pi(x) - f(x)| > c/3 \). \( \square \)

We remark that using a more modern result of Rankin [12] (or Erdős [1]), we have \( |\pi(x) - f(x)|/(\log \log x)^{1-e} \) unbounded for any \( e > 0 \).

3. On \( p_n^2 - p_{n-p_{n+i}} \). Let \( M(n) = \max_{0 < i < n} p_{n-p_{n+i}} \). By the Corollary to Theorem 2.2, there are infinitely many \( n \) with \( p_n^2 - M(n) > 0 \). We now show this difference can be arbitrarily large.

**Theorem 3.1.** \( \lim \sup (p_n^2 - M(n)) = \infty. \)

**Proof.** We shall show that \( p_n^2 - M(n) \to \infty \) as \( n \) runs through the set \( P'(1/2) \), the first coordinates of the vertex points on the boundary of the convex hull of \( \{(m, \sqrt{p_m}): m = 1, 2, \ldots \} \). Say \( n \in P'(1/2) \) and \( i < n \) is such that

\[
(3.1) \quad p_n - \sqrt{p_n} < p_{n-i}, \quad p_{n+i} < p_n + \sqrt{p_n}.
\]

Let \( d = p_n - p_{n-i} \), \( d' = p_{n+i} - p_n \). Then

\[
(3.2) \quad 0 < p_n^2 - M(n) \leq p_n^2 - p_{n-p_{n+i}} = (d - d')p_n + dd' < (d - d' + 1)p_n.
\]

Hence, we must have \( d' < d \). Thus, we have either (1) \( d' = d \) or (2) \( d' < d \). In the second case, we have from (3.2) that

\[
(3.3) \quad p_n^2 - p_{n-p_{n+i}} > (d - d')p_n \geq p_n.
\]

We shall now show that for each \( e > 0 \), there is an \( n_0(e) \) such that for all \( n > n_0(e), n \in P'(1/2) \), we have

\[
(3.4) \quad p_n - p_{n-1} > (1 - e)\log n.
\]

Assuming (3.4), we have for the above case (1) and large \( n \) that

\[
(3.5) \quad p_n^2 - p_{n-p_{n+i}} = dd' = d^2 > (1 - 2e)\log^2 n.
\]

To prove (3.4), let \( f \) be the function whose graph is the upper boundary of the convex hull of \( \{(n, \sqrt{p_n}): n = 1, 2, \ldots \} \). By the prime number theorem, \( f(x) \sim \sqrt{x} \log x \) as \( x \to \infty \). Since \( f \) is concave, we have \( f'(x) \sim \frac{1}{2}\sqrt{\log x/x} \) as \( x \to \infty \), ignoring the isolated values of \( x \) where \( f'(x) \) does not exist. (Let \( \delta > 0 \) be small. Since \( f \) is concave, we have...
THE PRIME NUMBER GRAPH

\[
\frac{f((1 + \delta)x) - f(x)}{\delta x} \leq f'(x) \leq \frac{f(x) - f((1 - \delta)x)}{\delta x}
\]
for all \(x\) where \(f'(x)\) exists. Multiplying the inequality by \(2\sqrt{x}/\log x\) and letting \(x \to \infty\), we have

\[
2(\sqrt{1 + \delta} - 1)/\delta \leq \lim_{x \to \infty} 2\sqrt{x}/\log x f'(x) \leq 2(1 - \sqrt{1 - \delta})/\delta.
\]

Now letting \(\delta \to 0\), we have \(\lim_{x \to \infty} 2\sqrt{x}/\log x f'(x) = 1\).}

The slope of the segment joining \((n, \sqrt{p_n})\) to \((n - 1, \sqrt{p_{n-1}})\) is

\[
\sqrt{p_n} - \sqrt{p_{n-1}} < \frac{p_n - p_{n-1}}{2\sqrt{p_n}}.
\]

Hence, if (3.4) fails, this segment has slope less than

\[
\frac{(1 - \epsilon)\log n}{2\sqrt{p_{n-1}}} < \frac{1 - \epsilon/2}{2} \sqrt{\frac{\log n}{n}} < f'(n - 0)
\]

for large \(n\). Thus, if \(n \in P'(1/2)\) and \(n\) is large, we have \(f(n) = \sqrt{p_n}\) so that \((n - 1, \sqrt{p_{n-1}})\) lies above \((n - 1, f(n - 1))\), a contradiction. Thus, we have established (3.4).

Now assume \(n \in P'(1/2)\), \(n\) is large, and \(i < n\) is such that (3.1) fails. Let \(p = p_{n-i}, q = p_{n+i}, D = q - p\). Then using \(P'(1/2) \subset P(1/2)\) (cf. (2.3)),

\[
p_n - \sqrt{pq} > ((\sqrt{p} + \sqrt{q})/2)^2 - \sqrt{pq} = (D/2(\sqrt{p} + \sqrt{q}))^2 > D^2/16q.
\]

If \(q < 2p_n\), then \(D^2/16q > 1/32\) since our assumption that (3.1) fails implies \(D > \sqrt{p_n}\). If \(q > 2p_n\), then \(D > q/2\), so \(D^2/16q > q/64\). Thus, for large \(n\) and either case,

\[
p_n - \sqrt{p_{n-i}p_{n+i}} > 1/32,
\]

so that

\[
(3.6) \quad p_n^2 - p_{n-i}p_{n+i} > p_n/32.
\]

Our theorem now follows from (3.3), (3.5), and (3.6). \(\square\)

We remark that we have actually proved

\[
\lim \sup (p_n^2 - M(n))/\log^2 n \geq 1.
\]

4. Collinear Points on the Prime Number Graph. A well-known conjecture is that for every \(k\) there are \(k\) primes which form an arithmetic progression. There is even a stronger conjecture that the \(k\) primes be consecutive. If this stronger conjecture were true, we would have the corollary that for every \(k\) there are \(k\) points on the prime number graph which are collinear. We now prove this result without any unproved hypothesis.

**Theorem 4.1.** For every \(k\) there are \(k\) collinear points on the prime number graph.
Proof. It is enough to show that for every sufficiently large \( k \), there are \( k \) colinear points on the inverse of the prime number graph:

\[ \{(p_n, n): n = 1, 2, \ldots \}. \]

We note that by the prime number theorem with error term, we have an \( x_0 \) such that for all \( x > x_0 \),

\[ |\pi(x) - \text{li}(x)| < x / \log^4 x. \tag{4.1} \]

Let \( k \) be a large enough integer so that if \( u = e^k \), \( v = u + u / \log u \), then

\[ u > x_0, \quad 2u / \log^4 u > v / \log^4 v. \]

Let \( T \) be the parallelogram bounded by the two vertical lines \( x = u, x = v \) and by the two lines with slope \( 1/k \) through

\[ (u, \text{li}(u) + 2u / \log^4 u), \quad (u, \text{li}(u) - 3u / \log^4 u). \]

We now note that if \( u < x < v \) and \( |y - \text{li}(x)| < x / \log^4 x \), then \((x, y)\) is in \( T \). This result is checked by an elementary calculation, the details of which we suppress.

We have by (4.1) and the above, that if \( u < p_n < v \), then \((p_n, n)\) lies in \( T \). Now consider lines passing through lattice points in the interior of \( T \) with slope \( 1/k \). There are exactly

\[ 5k u / \log^4 u = 5u / \log^3 u \]

such lines. But by (4.1),

\[ \pi(v) - \pi(u) \sim (v - u) / \log u = u / \log^2 u, \]

so that we have

\[ \pi(v) - \pi(u) > u / 2 \log^2 u. \]

Thus, one of the above-mentioned lines passes through at least

\[ (5u / \log^3 u)^{-1} (u / 2 \log^2 u) = k / 10 \]

points \((p_n, n)\) with \( u \leq p_n \leq v \). \( \Box \)

It was pointed out to me by P. Erdős that this proof generalizes to give the following striking result: for any given \( k \), the set of \( n \) for which \((n, p_n)\) is not on a line with \( k \) other points \((m, p_m)\) has density 0.

It is interesting that the above proof uses an error term in the prime number theorem. (Rather than (4.1), all that we really need to prove the theorem is an error term of \( o(x / \log^2 x) \).) If \( a_1 < a_2 < \ldots \) is a sequence of integers with \( a_n = (1 + o(1)) n \log n \), we cannot prove that for every \( k \) there are \( k \) points \((n, a_n)\) collinear, but we nonetheless conjecture that this is the case. We can prove that if \( f(x) \) is a smooth convex function with \( f(x) = O(x / \sqrt{\log x}) \) and \( a_n = (1 + o(1)) f(n) \), then for every \( k \) there are \( k \) points \((n, a_n)\) collinear.

In a forthcoming paper we shall show that if the increasing sequence of integers
$a_n$ does not have density 0, then for every $k$ there are $k$ of the $(n, a_n)$ collinear. More generally, we shall prove that given any $k$ and $B$ there is a number $n(k, B)$ such that if $n > n(k, B)$ and $u_0, u_1, \ldots, u_n$ is a sequence of points in $\mathbb{Z}^2$ with

$$\frac{1}{n} \sum_{i=1}^{n} \|u_i - u_{i-1}\| \leq B,$$

then $k$ of the $u_i$ are collinear. For related results see Ramsey [10] and Ramsey and Gerver [11].

5. Further Comments and Problems. Recalling from Section 2 the definition of $P(\alpha)$, we note that for every $\alpha \in [0, 1)$,

$$(5.1) \quad \bigcup_{\beta > \alpha} P(\beta) = P(\alpha).$$

Indeed, if $n \in P(\alpha)$, since the inequalities (2.3) and (1.1) are strict, it follows that there is an $\varepsilon = \varepsilon(n, \alpha) > 0$, such that $n \in P(\alpha + \varepsilon)$. It would be interesting to study the set

$$A: = \{ \alpha \in (0, 1): \bigcap_{\beta < \alpha} P(\beta) \neq P(\alpha) \}.$$  

Note that if $n \in P(0)$, then from (5.1) we have an $\alpha_n, 0 < \alpha_n \leq 1$, with

$$\{ \alpha: n \in P(\alpha) \} = [0, \alpha_n).$$

However,

$$(5.2) \quad \bigcap_{\alpha < 1} P(\alpha) = \{ 1 \},$$

so that if $n \in P(0), n > 1$, then $\alpha_n < 1$. To see (5.2), note that if $n \in \bigcap_{\alpha < 1} P(\alpha), n > 1$, then

$$(5.3) \quad 2p_n \geq p_1 + p_{2n-1}.$$  

However, from Landau's theorem that $2\pi(x) - \pi(2x) \to \infty$ as $x \to \infty$, (5.3) can hold but for at most finitely many $n > 1$. Recent work of L. Schoenfeld [14] can be used to show (5.3) in fact holds for no $n > 1$. We now easily see that

$$A = \{ \alpha_n: n \in P(0), n > 1 \}.$$  

Thus, $A$ is a countably infinite subset of $(0, 1)$. We conjecture that $A$ is dense in $[0, 1]$. An elementary Galois theory argument shown to me by J. Shyr gives that every member of $A$ is irrational. (Note that $\alpha_n \in A$ implies there is an $i < n$ with $2p_n^{\alpha_n} = p_n^{\alpha_{n-i}} + p_n^{\alpha_{n+i}}$.)

Recalling the definition of $M(n)$ from Section 3, we conjecture that

$$(5.4) \quad \limsup(p_n^2 - M(n))/p_n > 0.$$  

Note that the proof of Theorem 3.1 would give this conjecture if we could show that there are infinitely many $n \in \mathcal{P}'(1/2)$ such that $2p_n \neq p_{n-i} + p_{n+i}$ for all $i$ satisfying
If true, (5.4) is probably very close to best possible. Using Cramér’s conjecture

\[ \lim \sup \frac{(p_n - p_{n-1})/\log^2 n}{p_n} = 1, \]

we have as \( n \to \infty \)

\[ M(n) \geq p_{n+1}p_{n-1} \geq p_n^2 - (1 + o(1))p_n \log^2 n, \]

so that

\[ \lim \sup \frac{(p_n^2 - M(n))/p_n \log^2 n}{p_n} \leq 1. \]

Of course the prime number theorem gives \( |p_n^2 - M(n)| = o(p_n^2) \). In fact, using the best known error term in the prime number theorem gives

\[ |p_n^2 - M(n)| = O(p_n^2 \cdot \exp(-c(\log n)^{3/5})) \]

for some \( c > 0 \). If the Riemann hypothesis is assumed, we would have

\[ |p_n^2 - M(n)| = O(p_n^{3/2} \log^2 n). \]

We now briefly investigate the maximal order of \( M(n) - p_n^2 \). From Erdős [2] it follows that there is a \( c > 0 \) such that

\[ p_{n+1} - p_n > (1 + c)(p_n - p_{n-1}) \]

has a positive density of solutions \( n \). But from an easy application of Brun’s method, we have the upper density of the set of \( n \) for which \( p_n - p_{n-1} < t \log n \) tends to 0 as \( t \to 0 \). Thus, there is a \( t_0 > 0 \) such that there is a positive density of \( n \) for which (5.5) holds and

\[ p_n - p_{n-1} > t_0 \log n. \]

For these \( n \),

\[ p_{n+1}p_{n-1} - p_n^2 \geq (1 + o(1))ct_0p_n \log n, \]

so that

\[ \lim \sup \frac{(M(n) - p_n^2)/p_n \log n}{p_n} > 0. \]

We now define \( A(n) = \min_{0 < i < n}(p_{n-i} + p_{n+i}) \). Using another result from Erdős [2] and an argument similar to the above, we have

\[ \lim \sup \frac{(2p_n - A(n))/\log n}{p_n} > 0. \]

From the Corollary to Theorem 2.1 we have infinitely many \( n \) for which \( A(n) > 2p_n \). We conjecture that \( A(n) - 2p_n \) can be arbitrarily large. In a short computer search for nonnegative occurrences of \( A(n) - 2p_n \), kindly performed for us by E. R. Canfield, the largest value of \( A(n) - 2p_n \) found for \( n < 1000 \) was 24, assumed at \( n = 985 \) (\( p_{985} = 7759 \)). In addition, it was noticed that values of \( p_n \) for which \( A(n) - 2p_n \geq 0 \) are distributed very nearly like the squares. However, we cannot show such \( p_n \) have relative density 0 in the primes, although this surely must be the case. We also conjecture that the set of \( n \) for which \( p_n^2 > M(n) \) has density 0. More numerical work might prove profitable in examining these problems.

In this regard, we should point out that the set of \( n \) for which \( (n, p_n) \) is a vertex point on the boundary of the convex hull of the prime number graph has density 0. For if \( n_1 < n_2 < \ldots \) are these values of \( n \), then we easily see that \( p_{n_2/n_2} < p_{n_3/n_3} < \ldots \). We can then apply a result of Erdős and Prachar [5] (also see Rieger [13]) to get that the \( n_i \) have density 0. However, we cannot prove that the set of \( n \) for
which \((n, \log p_n)\) is a vertex point on the boundary of the convex hull of the set of all \((m, \log p_m)\) has density 0. A result of Erdős and Rényi [6] might prove helpful in this regard.

Combining the method of proof of Theorem 2.1 with the best known results on the error term of the prime number theorem, we can show there is a positive constant \(c\) such that for all sufficiently large \(x\), the number of \(n \leq x\) for which \((1.2)\) holds is at least \(\exp(c(\log x)^{3/5})\). Assuming the Riemann hypothesis would give at least \(c'x^{1/4}(\log x)^{-3/2}\) such \(n \leq x\). We get similar results for the distribution of the values of \(n\) for which \((1.1)\) holds. We conjectured above that both of these sets have density 0. We should remark that the results of Erdős [2] give that both sets have upper density strictly less than 1.

The Corollary to Theorem 2.2 has an immediate application to a certain extremum problem in combinatorial number theory. In the paper Erdős [3], \(f(k, x)\) is defined as the least integer \(r\) such that whenever \(1 < a_1 < \cdots < a_k \leq x\) are integers, there are \(r + 1\) \(a\)'s composed of at most \(r\) different primes. Then, if \(k \geq \pi(x) + 1\), \(f(k, x)\) exists. In [3] Erdős gives proofs due to him and Straus that

\[
f(k, x) \leq 2\pi(\sqrt{x}) + 1
\]

for all \(x\) and

\[
f(k, x) \sim 2\pi(\sqrt{x}).
\]

They conjecture

\[
2\pi(\sqrt{x}) - f(\pi(x) + 1, x) \to \infty.
\]

It is in this regard that they make the conjecture stated at the beginning of this paper.

We now show that \((5.7)\) is false by showing that equality can hold in \((5.6)\) for arbitrarily large values of \(x\). We conjecture the expression in \((5.7)\) is unbounded, but we cannot show this.

It is shown in [3] that \(f(\pi(x) + 1, x) \geq t + 1\), where \(t\) is the largest integer so that all products \(p_i p_j\) with \(i + j = t + 2\) and \(i \neq j\) do not exceed \(x\). Let now \(n\) be an integer for which \((1.1)\) holds. Let \(x = p_n^2 - 1\). Then the \("t"\) for this \(x\) is at least \(2n - 2\), so

\[
f(\pi(x) + 1, x) \geq 2n - 1 = 2\pi(\sqrt{x}) + 1.
\]

Since \((1.1)\) holds for infinitely many \(n\), we have thus shown equality holds in \((5.6)\) for an unbounded set of \(x\).


14. L. SCHOENFELD, (In preparation.)
