THE CONVEX HULL OF THE PRIME NUMBER GRAPH

NATHAN MCNEW

ABSTRACT. Let p_n denote the n-th prime number, and consider the prime number graph, the collection of points (n, p_n) in the plane. Pomerance uses the points lying on the boundary of the convex hull of this graph to show that there are infinitely many n such that $p_{2n} < p_{n-i} + p_{n+i}$ for all $i < n$. More recently, the primes on the boundary of this convex hull have been considered by Tutaj. We resolve several conjectures of Pomerance and Tutaj by giving improved bounds on the number and distribution of these primes as well as related forms of 'extreme' primes.

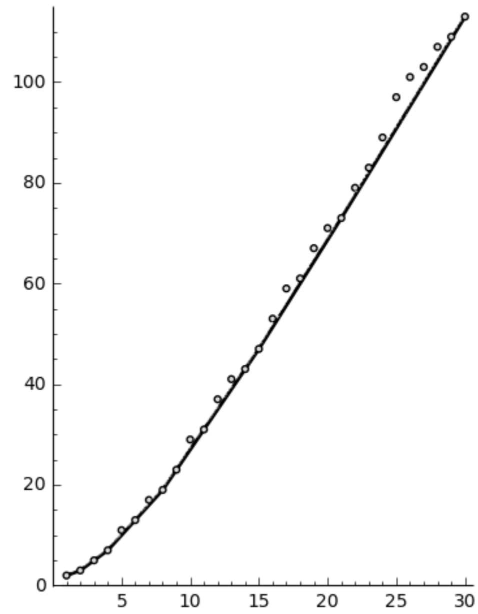
1. INTRODUCTION

Let p_n denote the n-th prime number, and consider the collection of points (n, p_n) in the plane $\mathbb{R}^2$, which we refer to as the *prime number graph*. The work of Zhang [12] and Maynard [6] and Tao implies there exist infinitely many pairs of consecutive points in this graph with some fixed, finite, integral slope between them. Currently we know this slope is at most 246, and the twin-prime conjecture is true if and only if there are infinitely many such pairs with slope exactly 2. On the other hand, from the prime number theorem, we know that this is not the behaviour of these points on average, as the collection of points which form the prime number graph tends upward faster than any linear function, growing roughly as $n \log n$, so the slopes between the first n consecutive points is about $\log n$ on average.

Because of the irregularities in the gaps between primes, the growth and distribution of these points can be fairly erratic. For example, if one considers the shape that is created by taking all of the line segments between consecutive points in this graph, the result is far from being a convex subset of the plane.

One can take the convex hull of this shape, however, which smooths out most of the irregularity in the primes' distribution, and ask about the collection of points (n, p_n), which are vertex points of this convex hull. The subset of primes forming such points, which we will refer to in the following as the *convex primes*, was studied by Pomerance in [9], and recently by Tutaj [11] and additionally is discussed in problem A14 of Guy's book of unsolved problems in number theory [5]. In what follows we will let $c_1, c_2, \ldots$ denote the indices of the sequence of convex primes, $p_{c_1}, p_{c_2}, \ldots$.

Figure 1. The first 30 points on the prime number graph, along with the lower boundary of the convex hull of the prime number graph.



Pomerance uses the set of convex primes, which he shows is an infinite set, to study a second subset of the primes, the *midpoint convex primes*, which are those prime numbers, p_n, which satisfy the inequality

$$2p_n < p_{n-i} + p_{n+i} \qquad \text{for all positive } i < n. \tag{1}$$

Because this condition is equivalent to the requirement that the line segment connecting any two points $(n-i, p_{n-i})$, and $(n+i, p_{n+i})$ pass above the point (n, p_n) we see that the set of convex primes is clearly a subset of the midpoint convex primes, and so the infinitude of the former set immediately implies the infinitude of the latter.

Pomerance also looks at a multiplicative version of inequality (1), those primes which satisfy the inequality

$$p_n^2 > p_{n-i}p_{n+i} \qquad \text{for all positive } i < n. \tag{2}$$

Primes satisfying this condition have become known as *good primes*. Using the log-prime number graph, the collection of points $(n, \log p_n)$, and specifically those primes, referred to here as the *log-convex primes*, corresponding to the vertices of the convex hull of the log-prime number graph, Pomerance likewise shows that there are infinitely many good primes. This disproved a conjecture of Erdős, who had conjectured that there were only finitely many good primes, while confirming one of Selfridge, who had conjectured the opposite.

Consider the line connecting the origin (0,0) with any vertex of the convex hull, (n, p_n). This line has slope $\frac{p_n}{n}$ and, with the exception of the first few convex primes (specifically the points (1,2), (2,3), and (4,7)), one finds that the portion of the line segment to the right of the line $x = 1$ lies entirely within this convex hull, hence strictly above and to the left of its lower boundary. This means that the slopes of these lines must be increasing, and so within the sequence of convex primes $p_{c_1}, p_{c_2}, \ldots$ we have necessarily that

$$\frac{p_{c_i}}{c_i} < \frac{p_{c_{i+1}}}{c_{i+1}} \tag{3}$$

for all $i \geq 2$.

Pomerance points out that a result of Erdős and Prachar [4], which shows that any subsequence of the primes which has the property (3) has relative density 0 in the primes, clearly implies the same for the convex primes, and so the count of the convex primes up to x is $o\left(\frac{x}{\log x}\right)$. Pomerance also notes that another result of Erdős [3] shows that the midpoint convex primes have relative upper density strictly less than one as well.

Note that while the convex primes are closely related to the set of primes p_n with the property that

$$\frac{p_n}{n} < \min_{1 \leq i < \infty} \frac{p_{n+i}}{n+i}, \tag{4}$$

they are not the same set, and in practice it appears that the set of primes satisfying (4) is a substantially larger set, of which the convex primes appear to be a subset.

Tutaj [11] proves the following theorem, conditional on the Riemann Hypothesis.

Theorem 1.1 (Tutaj). *Assume the Riemann hypothesis, and let p_{c_i} denote the i-th convex prime. Then*

$$\lim_{i \to \infty} \frac{p_{c_{i+1}}}{p_{c_i}} = 1.$$

He also makes several conjectures, two of which we restate here for reference.

Conjecture 1.2 (Tutaj). *The sum of the reciprocals of the convex primes,*

$$\sum_{i=1}^{\infty} \frac{1}{p_{c_i}}$$

converges.

Conjecture 1.3 (Tutaj). *The sum of the reciprocals of the logarithms of the convex primes,*

$$\sum_{i=1}^{\infty} \frac{1}{\log p_{c_i}}$$

diverges.

In Theorem 2.2 we give a substantially better upper bound than $o(\pi(x))$ for the convex primes, namely that their count is $O\left(\pi(x)^{2/3}\right)$. Conjecture 1.2 follows as a corollary. Both Conjecture 1.3 and an unconditional version of Theorem 1.1 follow as corollaries to Theorem 2.3, in which we prove an upper bound for the size of a the gap between two convex primes. (We also get a substantial improvement to this result, Theorem 2.4, if we assume the Riemann hypothesis.)

Additionally, we get a lower bound, $e^{\log^{3/5-\epsilon} x}$ for the number of convex primes up to x, and assuming the Riemann Hypothesis this can be improved to $\frac{b' x^{1/4}}{\log^{3/2} x}$ for some constant $b' > 0$. These results were claimed without proof in [9]. In

Section 4 we confirm another conjecture Pomerance made in that paper, that the log-convex primes have relative density zero in the primes.

Finally, in Section 5 we give results of computations on the convex primes and log-convex primes up to 10^{13}. Based on these computations it appears that the exponent $1/4$ arising in the lower bound of the convex primes is likely to be closer to the correct power of x in the counting function of the convex primes up to x.

2. Counting the Convex Primes

We give here a substantially improved upper bound for the count of the convex primes, using only the prime number theorem, and the fact that there aren't many possible rational slopes with small denominator. First however, we prove a lemma regarding the slope of a line segment along the edge of the convex hull.

Lemma 2.1. *If (m, p_m) is any point on the boundary of the convex hull of the prime number graph, then the slope of the line segment of the convex hull following this point has slope*

$$\log m + \log\log m + o(1)$$

as $m \to \infty$.

Proof. Suppose, for the sake of contradiction, that there exist arbitrarily large m where (m, p_m) is on the boundary of the convex hull and the slope is greater than $\log m + \log\log m + d$ for some $d > 0$.

Fix $\alpha > 1$ chosen so that $\log\alpha + \frac{1+d}{\alpha} - d \leq 1 - \epsilon$ for some $\epsilon > 0$. Then for sufficiently large m, at those values of m where the slope of the convex hull following p_m is greater than $\log m + \log\log m + d$, we have (using the fact, from the prime number theorem, that $p_m = m(\log m + \log\log m - 1 + o(1))$) that

$$\begin{aligned}
p_{\lceil \alpha m\rceil} &\geq p_m + (\alpha m - m)(\log m + \log\log m + d)\\
&= m(\log m + \log\log m - 1 + o(1)) + (\alpha m - m)(\log m + \log\log m + d)\\
&= \alpha m(\log\alpha m + \log\log\alpha m - \log\alpha - \frac{1+d}{\alpha} + d + o(1))\\
&\geq \alpha m(\log\alpha m + \log\log\alpha m - 1 + \epsilon + o(1))
\end{aligned}$$

as $m \to \infty$, which would contradict the prime number theorem.

Similarly, if we suppose there are arbitrarily large m where the slope is less than $\log m + \log\log m - d$ for some $d > 0$, we can fix a $0 < \beta < 1$ chosen so that $\log\beta + \frac{1-d}{\beta} + d \leq 1 - \epsilon$ for some $\epsilon > 0$. Then for sufficiently large m, where the slope of the convex hull following p_m is less than $\log m + \log\log m - d$, we have that

$$\begin{aligned}
p_{\lceil \beta m\rceil} &\geq p_m - (m - \beta m)(\log m + \log\log m - d)\\
&= m(\log m + \log\log m - 1 + o(1)) - (m - \beta m)(\log m + \log\log m - d)\\
&= \beta m(\log\beta m + \log\log\beta m - \log\beta - \frac{1-d}{\beta} - d + o(1))\\
&\geq \beta m(\log\beta m + \log\log\beta m - 1 + \epsilon + o(1))
\end{aligned}$$

as $m \to \infty$, again contradicting the prime number theorem. ∎

We now get an upper bound for the count of the convex primes.

Theorem 2.2. *The count of the convex primes up to x is $O\left(\frac{x^{2/3}}{\log^{2/3} x}\right)$.*

Proof. We count those convex primes in the dyadic interval $(\frac{1}{2}x, x]$. The slopes between the consecutive convex primes are necessarily strictly increasing rational numbers given by

$$\frac{p_{c_{j+1}} - p_{c_j}}{c_{j+1} - c_j}.$$

Because we are counting those $p_{c_j} \in \left(\frac{1}{2}x, x\right]$, we have $\frac{1}{2}\pi(x) < \pi\left(\frac{1}{2}x\right) < c_j \leq \pi(x)$. From Lemma 2.1 we know that the slope at each p_{c_j} is contained in some interval of length $\log 2 + o(1)$. If $c_j - c_{j-1} = k$, this leaves at most $k(\log 2 + o(1))$ possible values of $p_{c_j} - p_{c_{j-1}}$. Thus, for each integer k, there are at most $O(k)$ pairs $p_{c_j}, p_{c_{j+1}}$ in $(\frac{1}{2}x, x]$ with $c_j - c_{j-1} = k$, or $O(K^2)$ such convex primes which follow a gap between indices of consecutive convex primes at most K apart, a parameter to be chosen shortly.

The number of consecutive convex primes $p_{c_{j-1}}$ and p_{c_j} in the interval $\left(\frac{1}{2}x, x\right]$ with $c_j - c_{j-1} > K$ is at most $O\left(\frac{x}{K \log x}\right)$. Equating K^2 with $\frac{x}{K \log x}$ gives $K = (x/\log x)^{1/3}$ and thus that the total number of convex primes in the interval $(\frac{1}{2}x, x]$ is $O\left(\frac{x^{2/3}}{\log^{2/3} x}\right)$. The result then follows by summing dyadically. □

Igor Shparlinski points out that this result also follows from a result of George Andrews [1]. Andrews shows that any convex region of the plane with area A, bounded by line segments whose endpoints lie in $\mathbb{Z}^2$, has at most $O(A^{1/3})$ non-collinear vertices. In this case, the region bounded below by the convex hull of the prime number graph and above by the line segment connecting (1,2) to the vertex corresponding to the greatest convex prime less than x is a convex region with area $O\left(\frac{x^2}{\log x}\right)$.

As mentioned in the introduction, this proves Tutaj's Conjecture 1.2, showing that the sum of the reciprocals of the convex primes converges. We now give an upper bound for the size of the gap between consecutive convex primes, which will, as a corollary give us a lower bound for the count of the convex primes.

Theorem 2.3. *There exists a constant $B > 0$ such that for sufficiently large values of i the gap between consecutive convex primes p_{c_i} and $p_{c_{i+1}}$, is bounded above by*

$$p_{c_{i+1}} - p_{c_i} \leq p_{c_i} \exp\left\{\frac{-B \log^{3/5} p_{c_i}}{(\log\log p_{c_i})^{1/5}}\right\}.$$

Proof. The best known error term for the prime number theorem tells us that

$$c_i = \mathrm{li}(p_{c_i}) + O\left(p_{c_i} \exp\left\{\frac{-A \log^{3/5} p_{c_i}}{(\log\log p_{c_i})^{1/5}}\right\}\right) \tag{5}$$

for some constant $A > 0$. Thus, there exists a constant $D > 0$ such that all of the points of the prime number graph lie between the (implicitly defined as a function of x) curves

$$x = \mathrm{li}(y) - Dy \exp\left\{\frac{-A \log^{3/5} y}{(\log\log y)^{1/5}}\right\} \tag{6}$$

and

$$x = \mathrm{li}(y) + Dy \exp\left\{\frac{-A\log^{3/5} y}{(\log\log y)^{1/5}}\right\}. \tag{7}$$

We will find the maximal length of a line segment bounded between these two curves, and thus the maximal length of a line segment forming part of the boundary of the convex hull of the prime number graph.

Because both of these curves are convex and curved upward with unbounded derivative, we see that any line partially lying between these two curves with a sufficiently steep, positive slope will intersect the outer curve (7) twice. Since the slopes of the lines forming the boundary of the convex hull tend to infinity and the line segments necessarily exist between these two curves such a segmen must lie below and to the right of the inner curve (6). In fact the longest such a line segment could possibly be, given these constraints, would be if it began and ended at the boundary of the outer curve and was tangent to the inner curve in between.

Suppose then that a line segment contained between these two curves intersects (7) at points (x_1, y_1) and (x_2, y_2), $x_1 < x_2$. Let $\Delta = y_2 - y_1$ and let $(x_0, y_0) = \left(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2}\right)$ be the midpoint of this line.

In order for our line segment to remain to the right of (6), we need in particular that the midpoint of this line segment does, and so

$$\begin{aligned}\mathrm{li}\, y_0 - Dy_0 \exp\left\{\frac{-A\log^{3/5} y_0}{(\log\log y_0)^{1/5}}\right\} &< x_0 = \frac{x_1+x_2}{2} \\ = \frac{\mathrm{li}\, y_1}{2} + \frac{D}{2} y_1 \exp\left\{\frac{-A\log^{3/5} y_1}{(\log\log y_1)^{1/5}}\right\} &+ \frac{\mathrm{li}\, y_2}{2} + \frac{D}{2} y_2 \exp\left\{\frac{-A\log^{3/5} y_2}{(\log\log y_2)^{1/5}}\right\}.\end{aligned} \tag{8}$$

Using the Taylor series for li$y$, we have that

$$\mathrm{li}\,(y+t) = \mathrm{li}\, y + \frac{t}{\log y} - \frac{t^2}{2y\log^2 y} + O\left(\frac{t^3}{y^2\log^2 y}\right) \tag{9}$$

and so

$$\begin{aligned}\mathrm{li}\, y_0 - \frac{\mathrm{li}\, y_1 + \mathrm{li}\, y_2}{2} &= \mathrm{li}\, y_1 + \frac{\Delta}{2} - \frac{\mathrm{li}\, y_1 + \mathrm{li}\, y_1 + \Delta}{2} \\ &= \mathrm{li}\, y_1 + \frac{\Delta}{2\log y_1} - \frac{\Delta^2}{8y_1\log^2 y_1} - \mathrm{li}\, y_1 - \frac{\Delta}{2\log y_1} + \frac{\Delta^2}{4y_1\log^2 y_1} + O\left(\frac{\Delta^3}{y_1^2\log^2 y_1}\right) \\ &= \frac{\Delta^2}{8y_1\log^2 y_1} + O\left(\frac{\Delta^3}{y_1^2\log^2 y_1}\right).\end{aligned} \tag{10}$$

Using this in (9), we get that

$$\begin{aligned}
&\frac{\Delta^2}{8y_1\log^2 y_1}+O\left(\frac{\Delta^3}{y_1^2\log^2 y_1}\right)\\
&\qquad< D\left(y_0\exp\left\{\frac{-A\log^{3/5}y_0}{(\log\log y_0)^{1/5}}\right\}+\frac{y_1}{2}\exp\left\{\frac{-A\log^{3/5}y_1}{(\log\log y_1)^{1/5}}\right\}\right.\\
&\qquad\qquad\left.+\frac{y_2}{2}\exp\left\{\frac{-A\log^{3/5}y_2}{(\log\log y_2)^{1/5}}\right\}\right)\\
&\qquad< D\exp\left\{\frac{-A\log^{3/5}y_1}{(\log\log y_1)^{1/5}}\right\}\left(y_0+\frac{y_1}{2}+\frac{y_2}{2}\right)\\
&\qquad= D\exp\left\{\frac{-A\log^{3/5}y_1}{(\log\log y_1)^{1/5}}\right\}(2y_1+\Delta), \qquad (11)
\end{aligned}$$

so that

$$\Delta^2 < Dy_1\log^2 y_1\exp\left\{\frac{-A\log^{3/5}y_1}{(\log\log y_1)^{1/5}}\right\}(2y_1+\Delta)+O\left(\frac{\Delta^3}{y_1}\right), \qquad (12)$$

which means that

$$\Delta \ll y_1\log y_1\exp\left\{\frac{-A\log^{3/5}y_1}{2(\log\log y_1)^{1/5}}\right\}. \qquad (13)$$

This upper bound for Δ means that the gap between the convex primes p_{c_i} and $p_{c_{i+1}}$ is

$$\begin{aligned}
p_{c_{i+1}}-p_{c_i} &\ll p_{c_i}\log p_{c_i}\exp\left\{\frac{-A\log^{3/5}p_{c_i}}{2(\log\log p_{c_i})^{1/5}}\right\}\\
&\le p_{c_i}\exp\left\{\frac{-B\log^{3/5}p_{c_i}}{(\log\log p_{c_i})^{1/5}}\right\}.
\end{aligned}$$

for some constant $B>0$. □

Note that if we assume the Riemann hypothesis, we can use the stronger form of the prime number theorem, that

$$n = \operatorname{li} p_n + O\left(\sqrt{p_n}\log p_n\right),$$

and the same proof gives the following improvement.

Theorem 2.4. *Assume the Riemann hypothesis, then the gap between consecutive convex primes p_{c_i} and $p_{c_{i+1}}$, is bounded above by*

$$p_{c_{i+1}}-p_{c_i} \ll p_{c_i}^{3/4}\log^{3/2}p_{c_i}.$$

It was claimed without proof in [9] that the best known results on the error term in the prime number theorem imply that the count of the convex primes up to x is at least $e^{c\log^{3/5-\epsilon}x}$ for some $c>0$. This follows as an immediate corollary to Theorem 2.3.

Corollary 2.5. *There exists a constant $B > 0$ such that the count of the number of convex primes up to x is at least*

$$\exp\left\{\frac{B\log^{3/5} x}{(\log\log x)^{1/5}}\right\}.$$

Likewise, if one assumes the Riemann Hypothesis, then we get the following as a corollary to Theorem 2.4.

Corollary 2.6. *Assume the Riemann Hypothesis. Then there exists a constant $B' > 0$ such that the count of the number of convex primes up to x is at least*

$$\frac{B' x^{1/4}}{\log^{3/2} x}.$$

The lower bound of Corollary 2.5 also has another corollary which proves Conjecture 1.3.

Corollary 2.7. *The sum of the reciprocals of the logarithms of the convex primes,*

$$\sum_{i=1}^{\infty} \frac{1}{\log p_{c_i}}$$

diverges.

Furthermore, by showing that $p_{c_{i+1}} - p_{c_i} = o(p_{c_i})$, Theorem 2.3 gives an unconditional proof of Theorem 1.1, that the ratio of consecutive convex primes converges to 1.

Corollary 2.8. *Let p_{c_i} denote the i-th convex prime. Then*

$$\lim_{i\to\infty} \frac{p_{c_{i+1}}}{p_{c_i}} = 1.$$

3. Edge convex primes

The results discussed above, except for Lemma 2.1, apply only for points (n, p_n) which are vertex points of the convex hull of the prime number graph. It is possible for points to lie on the boundary of the convex hull without being vertex points, for example, the point $(3, 5)$ lies on the line segment between the vertex points $(2, 3)$ and $(4, 7)$.

In the computation of the convex primes up to 10^{13} (discussed further in Section 5) we find only 5 examples of primes with this property, namely $5, 13, 23, 31$, and 43. Based on this evidence we conjecture the following.

Conjecture 3.1. *There are only finitely many primes p_n for which the point (n, p_n) lies on the boundary of the convex hull of the prime number graph without being a vertex point of it.*

Despite this conjecture, obtaining upper bounds for the number of these *edge convex primes* is more difficult than for the convex primes. Combining the results of Theorem 2.3 with the ideas of Theorem 2.2 we can get the following upper bound.

Theorem 3.2. *The count of the edge convex primes, those primes which lie on the boundary of the convex hull of the prime number graph without being vertices of it, is* $O\left(x\exp\left\{-b'\frac{\log^{3/5}x}{(\log\log x)^{1/5}}\right\}\right)$ *for some constant* $b'>0$.

Proof. We count those edge convex primes in the dyadic interval $(\frac{1}{2}x, x]$. As noted before, the slopes of the lines forming the boundary of the convex hull are necessarily strictly increasing rational numbers contained in some interval of length $\log 2+o(1)$, from Lemma 2.1. From Theorem 2.3 we see that there are necessarily at least $\exp\left\{B\frac{\log^{3/5}x}{(\log\log x)^{1/5}}\right\}$ distinct line segments in this interval (for some $B>0$), and that each line segment ranges over $O\left(x\exp\left\{-B\frac{\log^{3/5}x}{(\log\log x)^{1/5}}\right\}\right)$ primes.

If a line segment of the boundary of the convex hull has slope $\frac{a}{d}$, with a and d coprime, then the edge convex primes lying on this line segment must be spaced at least d primes apart, so the maximum number of edge convex primes contained on such an interval is $O\left(\frac{x}{d}\exp\left\{-B\frac{\log^{3/5}x}{(\log\log x)^{1/5}}\right\}\right)$

Since each of the slopes is a rational number contained in an interval of length $\log 2+o(1)$ there are $O\left(\varphi(d)\right)$ line segments forming the convex hull in this range with slope whose denominator in lowest terms is d, and thus the total number of edge convex primes that can lie on any line segment whose slope has denominator at most D is $O\left(Dx\exp\left\{-B\frac{\log^{3/5}x}{(\log\log x)^{1/5}}\right\}\right)$. Because the total number of edge convex primes which lie on those line segments whose slope has denominator (in reduced form) greater than D is $O\left(\frac{x}{D}\right)$, we find, by optimizing D, that the total number of edge convex primes in this interval is $O\left(x\exp\left\{-\frac{1}{2}B\frac{\log^{3/5}x}{(\log\log x)^{1/5}}\right\}\right)$. The result then follows by summing dyadically. □

Assuming the Riemann Hypothesis, this can be improved using Theorem 2.4.

Theorem 3.3. *Assuming the Riemann Hypothesis, we have that the count of the edge convex primes is* $O\left(x^{7/8}\log^{3/4}x\right)$.

4. Multiplicatively Convex Primes

In the paper [9] where Pomerance introduced the convex primes, he also considered the log-convex primes, those primes forming the vertices of the boundary of the convex hull of the log-prime number graph. Unlike the convex primes, he was not able to show that they had relative density zero among the primes, but conjectured that this was the case. We prove that this is in fact the case here. First however we prove a lemma regarding the slope of the boundary of the convex hull of the log-prime number graph.

In the following we adopt the notation of [2] and use $\mathrm{ali}(x)$ to denote the functional inverse of the logarithmic integral function, $\mathrm{li}(x)$.

Lemma 4.1. *If* $(m, \log p_m)$ *is any point on the boundary of the convex hull of the log-prime number graph, then the slope of the line segment of the convex hull following this point has slope*

$$\frac{\mathrm{ali}'(m)}{\mathrm{ali}(m)}+O\left(\frac{1}{m}\exp\left\{-\frac{A\log^{3/5}m}{2(\log\log m)^{1/5}}\right\}\right) \tag{14}$$

as $m \to \infty$, for some $A > 0$.

Proof. We can rewrite equation (5), the strongest known form of the prime number theorem, as

$$p_n = \operatorname{ali}(n) + O\left(n \exp\left\{-A\frac{\log^{3/5} n}{(\log\log n)^{1/5}}\right\}\right), \tag{15}$$

and so,

$$\log p_n = \log \operatorname{ali}(n) + O\left(\exp\left\{-A\frac{\log^{3/5} n}{(\log\log n)^{1/5}}\right\}\right). \tag{16}$$

Note that

$$\frac{d}{dx}\log \operatorname{ali}(x) = \frac{\operatorname{ali}'(x)}{\operatorname{ali}(x)} \sim \frac{\log x}{x\log x} = \frac{1}{x}$$

and that

$$\frac{d^2}{dx^2}\log \operatorname{ali}(x) \sim -\frac{1}{x^2}.$$

Suppose rather, for contradiction, that for any arbitrarily large constant C, there exist infinitely many d_i where the slope following the point $(d_i, \log p_{d_i})$ is less than $\frac{\operatorname{ali}'(d_i)}{\operatorname{ali}(d_i)} - \frac{C}{d_i}\exp\left\{-\frac{A\log^{3/5} d_i}{2(\log\log d_i)^{1/5}}\right\}$.

Let $f(x) = \exp\left\{-\frac{A\log^{3/5} x}{2(\log\log x)^{1/5}}\right\}$. Because the slopes of the boundary of the convex hull of the log-prime number graph are strictly decreasing, we see that for such values of d_i we have, using the Taylor series approximation of $\log \operatorname{ali}(x)$, that

$$\begin{aligned}
\log p_{\lfloor d_i + d_i f(d_i)\rfloor} &\le \log p_{d_i} + d_i f(d_i)\left(\frac{\operatorname{ali}'(d_i)}{\operatorname{ali}(d_i)} - \frac{C}{d_i}f(d_i)\right)\\
&= \log \operatorname{ali}(d_i) + d_i f(d_i)\frac{\operatorname{ali}'(d_i)}{\operatorname{ali}(d_i} - Cf(d_i)^2 + O\left(f(d_i)^2\right)\\
&= \log \operatorname{ali}\left(d_i + d_i f(d_i)\right) - Cf(d_i)^2 + O\left(f(d_i)^2\right).
\end{aligned} \tag{17}$$

Because the constant C can be taken arbitrarily large, equation (17) contradicts equation (16), thus proving that the slope is at least as large as expression (14). The corresponding upper bound follows by essentially the same argument. $\square$

We are now able to see that the log-convex primes have relative density zero among the primes.

Theorem 4.2. *The count of the number of log-convex primes up to x, as $x \to \infty$, is at most $\frac{x}{\log^{4/3-o(1)} x}$.*

Proof. Denote by $D(t, x)$ the number of primes $p \le x$ for which $p+t$ is also prime. Brun's sieve can be used to show (see for example [10]) that

$$D(t,x) \le c\left(\prod_{\substack{p\mid t\\ p\le x}}\left(1-\frac{1}{p}\right)^{-1}\right)\frac{x}{\log^2 x}$$

for some absolute constant c.

Fix x and consider pairs of primes

$$\frac{x}{\log x} \le p_n < p_{n+k} \le x \tag{18}$$

with $k \le \log^{1/3} x$. (The number of log-convex primes less than $\frac{x}{\log x}$ is already at most $\pi\left(\frac{x}{\log x}\right) \sim \frac{x}{\log^2 x}$ so their contribution is insignificant.)

Suppose a pair of primes p_n and p_{n+k} in this range are consecutive log-convex primes. Then the slope of the line conecting the corresponding points is $\frac{\log p_{n+k} - \log p_n}{k}$, and so Lemma 4.1 tells us that

$$\left| \frac{\log p_{n+k} - \log p_n}{k} - \frac{\operatorname{ali}'(n)}{\operatorname{ali} n} \right| = O\left(\frac{1}{n} \exp\left\{ -A \frac{\log^{3/5} n}{(\log\log n)^{1/5}} \right\} \right)$$

for some $A > 0$. Letting $\Delta = p_{n+k} - p_n$, we can write

$$\frac{\log p_{n+k} - \log p_n}{k} = \frac{\log(p_n + \Delta) - \log p_n}{k} = \frac{\Delta}{kp_n} + O\left(\frac{\Delta^2}{kp_n^2} \right). \tag{19}$$

So, assuming $\Delta \le n^\epsilon$, and using the fact that

$$p_n \frac{\operatorname{ali}'(n)}{\operatorname{ali} n} = \log n + O(\log\log n) = \log x + O(\log\log x)$$

for n in the range (18), we find that

$$|\Delta - k \log x| \le kb \log\log x \tag{20}$$

for some absolute constant b.

Thus, the total number of such pairs is bounded above by

$$\sum_{1 \le k \le \log^{1/3} x} \sum_{t \in \mathcal{I}_k} D(t, x) \le \sum_{1 \le k \le \log^{1/3} x} \sum_{t \in \mathcal{I}_k} c \left(\prod_{p \mid t} \left(1 - \frac{1}{p} \right)^{-1} \right) \frac{x}{\log^2 x} \tag{21}$$

where $\mathcal{I}_k = \{t : |t - k \log x| \le bk \log\log x\}$. Now, using Mertens' theorem,

$$\begin{aligned} \sum_{t \in \mathcal{I}_k} \prod_{p \mid t} \left(1 - \frac{1}{p} \right)^{-1} &< 2bk \log\log x \prod_{p \le k(\log x + b \log\log x)} \left(1 - \tfrac{1}{p} \right)^{-1} \\ &= O(k \log\log x \log(k \log x)) \\ &= O(k (\log\log x)^2). \end{aligned}$$

So,

$$\sum_{1 \le k \le \log^{1/3} x} \sum_{t \in \mathcal{I}_k} D(t, x) \ll \sum_{1 \le k \le \log^{1/3} x} k (\log\log x)^2 \frac{x}{\log^2 x} < (\log\log x)^2 \frac{x}{\log^{4/3} x}. \tag{22}$$

Any log-convex primes not involved in one of these pairs must be spaced at least $\log^{1/3} x$ primes distant from the next log-convex prime, so the count of such isolated log-convex primes is less than $\frac{1}{\log^{1/3} x} \pi(x)$. Thus the contribution from (22) dominates, and we see that the count of the log-convex primes up to x is at most $\frac{x}{\log^{4/3 - o(1)} x}$. □

5. Data and Future Work

The table below gives the count $C(x)$ of the number of convex primes up to x for x ranging up to 10^{13}, as well as the exponent one could raise x to in order to approximate this value of $C(x)$. Based on the data, it appears that the lower bound on the order of $\frac{x^{1/4}}{\log^{3/2} x}$, (obtained by assuming the Riemann Hypothesis, Corollary 2.6) may not be far from the truth, though it seems that $C(x)$ may be growing like x^c for some constant c closer to 0.285.

x	$C(x) = \#$ of convex primes up to x	$\frac{\log C(x)}{\log x}$
10^1	3	0.47712
10^2	6	0.38908
10^3	12	0.35973
10^4	22	0.33561
10^5	36	0.31126
10^6	65	0.30215
10^7	121	0.29754
10^8	223	0.29354
10^9	413	0.29066
10^{10}	756	0.28785
10^{11}	1409	0.28626
10^{12}	2621	0.28487
10^{13}	5150	0.28552

As mentioned in Section 3, there were only 5 edge convex primes in this range, all at most 43. From this data it seems natural to conjecture that there are only finitely many edge convex primes, even though the upper bound we have for their count, Theorem 3.2 is far weaker than our bound for the convex primes, Theorem 2.2.

Pomerance originally studied the convex primes as a way to show that there were infinitely many midpoint convex primes, those primes satisfying (3), since the convex primes are a subset of the midpoint convex primes. The midpoint convex primes appear to be far more numerous than the convex primes.

x	$M(x) = \#$ of midpoint convex primes up to x	$\frac{\log M(x)}{\log x}$
10^1	3	0.47712
10^2	8	0.45154
10^3	25	0.46598
10^4	89	0.48735
10^5	288	0.49188
10^6	1148	0.50999
10^7	4504	0.52194
10^8	17293	0.52973
10^9	71804	0.53957
10^{10}	283737	0.54529
10^{11}	1195764	0.55251

On the other hand, it seems clear from the data that the midpoint convex primes are a relatively sparse subset of the primes, growing slightly faster than $\sqrt{x}$, however it is still open to prove that their count is $o(\pi(x))$.

In light of (1), the midpoint convex primes can be characterized as those primes p_n, where the quantity

$$M_n = \min_{1\leq i<n} (p_{n+i} + p_{n-i}) - 2p_n \tag{23}$$

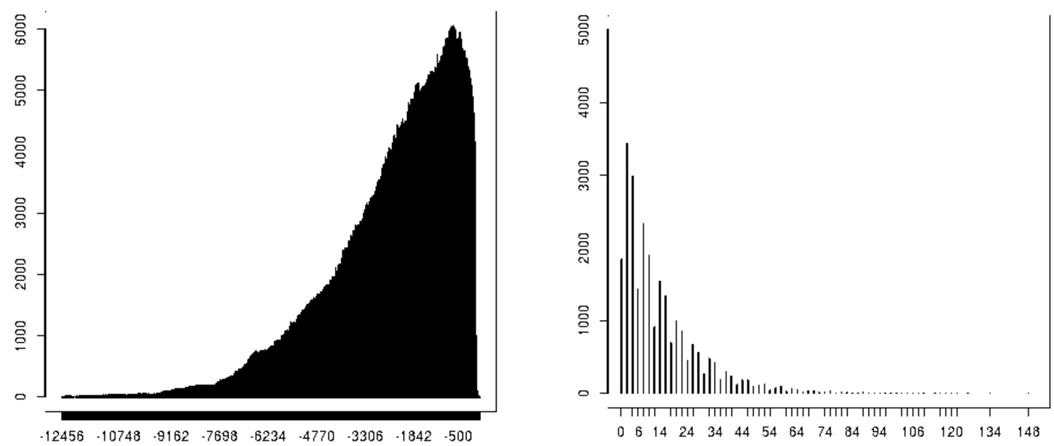
is positive. Included below are two histograms, showing the distribution of M_n for $n < 1.6 \times 10^8$. The first shows the entire distribution, the second shows the (very minor) tail of the distribution for non-negative values.

The distribution of this quantity M_n would be interesting to study further. Based on the data it appears likely that M_n can be arbitrarily large. It can easily be seen that M_n can be arbitrarily negative, as even

$$p_{n+1} + p_{n-1} - 2p_n = (p_{n+1} - p_n) - (p_n - p_{n-1}),$$

the difference between consecutive gaps, can be arbitrarily negative. In fact Pintz [8] shows that $\liminf\limits_{n\to\infty} \frac{p_{n+1}-p_n}{p_{n-1}-p_n} = 0$. Note also how the values of M_n have a tendency to avoid multiples of 6.

FIGURE 2. Distribution of the minimum of M_n for $n < 1.6 \times 10^8$ and the non-negative part of that distribution.



We also give the corresponding counts for the log-convex primes and the good primes, which appear to be somewhat more numerous than the convex primes and midpoint convex primes respectively.

x	$L(x)$ =# of log-convex primes up to x	$\frac{\log L(x)}{\log x}$	$G(x)$ = # of good primes up to x	$\frac{\log G(x)}{\log x}$
10^1	1	0	1	0
10^2	9	0.47712	11	0.52070
10^3	25	0.46598	44	0.54782
10^4	56	0.43705	176	0.56138
10^5	111	0.40906	671	0.56534
10^6	248	0.39908	2668	0.57103
10^7	533	0.38953	10942	0.57701
10^8	1060	0.37816	45150	0.58183
10^9	2182	0.37098	189365	0.58637
10^{10}	4555	0.36585		
10^{11}	9394	0.36117		
10^{12}	19510	0.35752		
10^{13}	40901	0.35475		

It would be interesting to develop a heuristic argument, possibly using a probabilistic model for the prime numbers, in order to conjecture what the correct asymptotic growth rate of these counts might be.

Obviously proving proving asymptotic formulae for these counts would be ideal, though this may be difficult. It would be interesting if progress could be made on even the following questions:

Question 5.1. *Can one prove that the counting function of the midpoint primes, $M(x)$ is $o(\pi(x))$? Or, likewise, that the counting function for the good primes $G(x)$ is $o(\pi(x))$?*

Question 5.2. *Clearly $C(x) < M(x)$ and $L(x) < G(x)$, can we prove that $C(x) = o(M(x))$, or that $L(x) = o(G(x))$?*

One could also consider the convex hull of other graphs related to the prime numbers, for example the collection of points $(\mathrm{li}(p_n), p_n)$, will have infinitely many points on its convex hull both above and below as the slopes of the bounding convex hull approach 1.

6. Acknowledgements

A portion of this work appeared in the author's Ph.D. thesis [7], written under the direction of Carl Pomerance, whose helpful suggestions were invaluable in the development of this paper. The author is also grateful to Igor Shparlinski, Robin Pemantle and the anonymous referee for useful suggestions.

References

1. G. Andrews, *A lower bound for the volume of strictly convex bodies with many boundary lattice points*, Trans. Amer. Math. Soc. **106** (1963), 270–279. MR 0143105 (26 #670)
2. J. Arias de Reyna and J. Toulisse, *The n-th prime asymptotically*, J. Théor. Nombres Bordeaux **25** (2013), no. 3, 521–555. MR 3179675
3. P. Erdős, *On the difference of consecutive primes*, Bull. Amer. Math. Soc. **54** (1948), 885–889. MR 0027009 (10,235b)

4. P. Erdős and K. Prachar, *Sätze und Probleme über p_k/k*, Abh. Math. Sem. Univ. Hamburg **25** (1961/1962), 251–256. MR 0140481 (25 #3901)
5. R. Guy, *Unsolved problems in number theory*, Springer, New York, 2004.
6. J. Maynard, *Small gaps between primes*, Ann. of Math. (2) **181** (2015), no. 1, 383–413. MR 3272929
7. N. McNew, *Multiplicative problems in combinatorial number theory*, Ph.D. thesis, Dartmouth College, 2015.
8. J. Pintz, *Polignac numbers, conjectures of Erdős on gaps between primes, arithmetic progressions in primes, and the bounded gap conjecture*, From arithmetic to zeta-functions, Springer, [Cham], 2016, pp. 367–384. MR 3644045
9. C. Pomerance, *The prime number graph*, Math. Comp. **33** (1979), no. 145, 399–408.
10. C. Pomerance and A. Sárközy, *Combinatorial number theory*, Handbook of combinatorics, Vol. 1, 2, Elsevier, Amsterdam, 1995, pp. 967–1018. MR 1373676 (97e:11032)
11. E. Tutaj, *Prime numbers with a certain extremal type property*, arXiv preprint arXiv:1408.3609 (2014).
12. Y. Zhang, *Bounded gaps between primes*, Ann. of Math. (2) **179** (2014), no. 3, 1121–1174. MR 3171761

Department of Mathematics, Towson University, Towson, MD 21286 USA