PRIMES BETWEEN CONSECUTIVE SQUARES

D. BAZZANELLA

Abstract

A well known conjecture about the distribution of primes asserts that between two consecutive squares there is always at least one prime number. The proof of this conjecture is quite out of reach at present, even under the assumption of the Riemann Hypothesis. The aim of this paper is to provide a bound for the exceptional set for this conjecture, unconditionally and under the assumption of some classical hypothesis. We also provide a conditional proof of the conjecture assuming an hypothesis about the behavior of Selberg's integral in short intervals.

Mathematical Subject Classification :
11NO5 - Distribution of Primes

This is the authors' post-print version of an article published on
Arch. Math. (Basel) 75 (2000) n.1, 29–34,
DOI:10.1007/s000130050469.[1]

1. - INTRODUCTION.

A well known conjecture about the distribution of primes asserts that

Conjecture 1.
For all integer n the interval $[n^2, (n+1)^2]$ contains a prime.

The proof of this conjecture is quite out of reach at present, even under the assumption of the Riemann Hypothesis (RH).

At any rate it is not difficult to prove unconditionally that Conjecture 1 holds for almost all integers n and, more precisely, we can prove that in almost all intervals $[n^2, (n+1)^2]$ there is the expected number of primes.

To have some results in the direction of Conjecture 1 we need to assume hypothesis stronger RH.

Define

$$J(N,h)=\int_1^N(\theta(x+h)-\theta(x)-h)^2dx,$$

with $\theta(x)=\sum_{p\le x}\log p$ and p prime number, and consider the following strong form of Montgomery's pair correlation conjecture

Conjecture 2.

$$J(N,h)=hN\log(N/h)-(\gamma+\log 2\pi)hN+o(hN)+O(N),$$

uniformly for $1\le h\le N^{1-\varepsilon}$.

Goldston [4] deduced the validity of Conjecture 1 assuming Conjecture 2.

In this paper we investigate about the exceptional set for the distribution of primes between two consecutive primes unconditionally and assuming some classical hypothesis.

The basic idea of this paper is to connect the exceptional set for the distribution of primes in intervals of type $[n^2,(n+1)^2]$ to the exceptional set for the asymptotic formula of the distribution of primes in short intervals, and using the properties of this set, see Bazzanella [1] and Bazzanella and Perelli [2], to obtain the desired results.

Our main unconditional result is the following

Theorem 1 *Let $\varepsilon>0$ then for every $[n^2,(n+1)^2]\subset[1,N]$ with $O(N^{1/4+\varepsilon})$ exceptions we have the expected number of primes.*

An immediate consequence of Theorem 1 is the following

Corollary 1 *Let $\varepsilon>0$ then for every $n\le N$ with $O(N^{1/4+\varepsilon})$ exceptions the interval $[n^2,(n+1)^2]$ contains a prime.*

Under RH we obtain in a similar way the following

Theorem 2 *Assume RH and let $f(x)\to\infty$ arbitrarily slowly. Then for every $[n^2,(n+1)^2]\subset[1,N]$ with $O(f(N)\log^2 N)$ exceptions we have the expected number of primes.*

Assuming hypothesis stronger than RH we can obtain a smaller exceptional set. In order to obtain a conditional proof of Conjecture 1 we are mainly interested to find the minimal unproved hypothesis to have an empty exceptional set. With this in mind we state the following conjecture

Conjecture 3.

$$J(N+Y,h) - J(N,h) = o(hN),$$

uniformly for $1 \le Y \le N^{1/2}$ and $N^{1/2} \ll h \ll N^{1/2}$.

Let $\psi(x) = \sum_{n \le x} \Lambda(n)$, where $\Lambda(n)$ is the von Mangoldt function, and assuming this conjecture we can obtain the last theorem

Theorem 3 *Assume Conjecture 3. Then we have*

$$\psi((n+1)^2) - \psi(n^2) \sim 2n \quad as \quad n \to \infty. \tag{1}$$

We note that Conjecture 3 is weaker than Conjecture 2, nevertheless Theorem 3 is stronger than result of Goldston [4] which asserts only the existence of a prime in intervals of type $[n^2, (n+1)^2]$.

2. - PRELIMINARY LEMMAS.

As in author's paper with A. Perelli [2] we define a set concerning the asymptotic formula

$$\psi(x + h(x)) - \psi(x) \sim h(x) \quad \text{as} \quad x \to \infty \tag{2}$$

as

$$E_\delta(X,h) = \{X \le x \le 2X : |\Delta(x,h)| \ge \delta h(x)\},$$

with $h(x)$ increasing function such that $x^\varepsilon \le h(x) \le x$ for some $\varepsilon > 0$ and

$$\Delta(x,h) = \psi(x + h(x)) - \psi(x) - h(x).$$

It is clear that (2) holds if and only if for every $\delta > 0$ there exists $X_0(\delta)$ such that $E_\delta(X,h) = \emptyset$ for $X \ge X_0(\delta)$. Hence for small $\delta > 0$, X tending to ∞ and $h(x)$ suitably small with respect to x, the set $E_\delta(X,h)$ contains the exceptions, if any, to the expected asymptotic formula for the number of primes in short intervals.

Moreover we consider the case $h(x) = 2\sqrt{x} + 1$ and introduce the functions

$$\mu_\delta = \inf\{\xi \ge 0 : |E_\delta(N, h(x))| \ll_{\delta,\theta} X^\xi\}$$

and

$$\mu = \sup_{\delta > 0} \mu_\delta, \tag{3}$$

to deal with the problem to estimate the exceptional set. For the proof of the main unconditional result we need the following estimate for the exceptional set for asymptotic formula (2), with $h(x) = 2\sqrt{x} + 1$.

Lemma 1 *Let μ defined by (3). Then we have*

$$\mu \leq \frac{3}{4}.$$

Proof.

We first reduce our problem to a similar one, but technically simpler. We begin by observing that if

$$|\{X \leq x \leq 2X : |\Delta(x, h(x))| \geq \frac{4\sqrt{X}}{\log X}\}| \ll X^{\alpha+\varepsilon} \tag{4}$$

holds with some $\alpha \geq 0$ and every $\varepsilon > 0$, then clearly $\mu \leq \alpha$. Further, given any $\varepsilon > 0$, we subdivide $[X, 2X]$ into $\ll X^{\varepsilon}$ intervals of the type $I_j = [X_j, X_j + Y]$ with $X \ll X_j \ll X$ and $Y \ll X^{1-\varepsilon}$. Writing $\xi_j = 2X_j^{-1/2}$ we have

$$\max_{x \in I_j} |h(x) - \xi_j x| \ll X^{1/2-\varepsilon}$$

uniformly in j, and hence

$$\Delta(x, h(x)) - (\psi(x + \xi_j x) - \psi(x) - \xi_j x) \ll X^{1/2-\varepsilon} \tag{5}$$

uniformly in j and $x \in I_j$.

From (4) and (5) is not difficult to see that if for some $\alpha \geq 0$ and any $\varepsilon > 0$

$$|\{X \leq x \leq 2X : |\psi(x + \xi_j x) - \psi(x) - \xi_j x| \geq \frac{2\sqrt{X}}{\log X}\}| \ll X^{\alpha+\varepsilon} \tag{6}$$

holds uniformly in j, then $\mu \leq \alpha$. Also, it is clear that in order to prove (6) we may restrict ourselves to the case $\xi_j = \xi = 2X^{-1/2}$, the other cases being completely similar.

In order to prove (6) we use the classical explicit formula, see ch. 17 of Davenport [3], to write

$$\psi(x + \xi x) - \psi(x) - \xi x = \sum_{|\gamma| \leq T} x^{\varrho} c_{\varrho}(\xi) + O(\frac{X \log^2 X}{T}), \tag{7}$$

uniformly for $X \leq x \leq 2X$, where $10 \leq T \leq X$, $\varrho = \beta + i\gamma$ runs over the non-trivial zeros of $\zeta(s)$,

$$c_{\varrho}(\xi) = \frac{(1+\xi)^{\varrho} - 1}{\varrho} \qquad \text{and} \qquad c_{\varrho}(\xi) \ll \min(X^{-1/2}, \frac{1}{|\gamma|}). \tag{8}$$

Choose

$$T = X^{1/2} \log^4 X, \tag{9}$$

and use the Jutila [8] density estimate, which asserts that for every $\varepsilon > 0$ we have

$$N(\sigma, T) \ll T^{2(1-\sigma)+\varepsilon} \qquad \text{for} \qquad \frac{11}{14} \leq \sigma \leq 1. \tag{10}$$

From (8), (9) and (10) by a standard argument we see that

$$\sum_{\substack{|\gamma|\leq T \\ \beta>11/14+\delta}} x^{\varrho} c_{\varrho}(\xi) \ll X^{-1/2} \log^2 X \max_{\sigma>11/14+\delta} X^{\sigma} N(\sigma, T) \ll \frac{X^{1/2}}{\log^2 X}, \tag{11}$$

for every $\delta > 0$ and uniformly for $X \leq x \leq 2X$.

Using Lemma 1 of Heath-Brown [6] we get

$$\int_X^{2X} | \sum_{\substack{|\gamma|\leq T \\ \beta\leq 11/14+\delta}} x^{\varrho} c_{\varrho}(\xi)|^4 dx \ll X^{-1+\varepsilon} \max_{\sigma\leq 11/14+\delta} X^{4\sigma} N^*(\sigma, T), \tag{12}$$

and by Theorem 2 of Heath-Brown [7] we have

$$N^*(\sigma, T) \ll \begin{cases} T^{(10-11\sigma)/(2-\sigma)+\varepsilon} & \dfrac{1}{2} \leq \sigma \leq \dfrac{2}{3} \\ T^{(18-19\sigma)/(4-2\sigma)+\varepsilon} & \dfrac{2}{3} \leq \sigma \leq \dfrac{3}{4} \\ T^{12(1-\sigma)/(4\sigma-1)+\varepsilon} & \dfrac{3}{4} \leq \sigma \leq 1. \end{cases} \tag{13}$$

Hence from (9), (12) and (13) we obtain

$$\int_X^{2X} | \sum_{\substack{|\gamma|\leq T \\ \beta\leq 11/14+\delta}} x^{\varrho} c_{\varrho}(\xi)|^4 dx \ll X^{11/4+\varepsilon},$$

for every $\varepsilon > 0$ and δ sufficiently small positive constant.

Hence we obtain

$$|\{X \leq x \leq 2X : | \sum_{\substack{|\gamma|\leq T \\ \beta\leq 11/14+\delta}} x^{\varrho} c_{\varrho}(\xi)| \geq \frac{\sqrt{X}}{\log X}\}| \ll X^{3/4+\varepsilon},$$

and then Lemma 1 is proved.

If we assume RH we have

Lemma 2 *Assume RH. Let* $h(x) = 2\sqrt{x}+1$ *and* $g(x) \to \infty$ *arbitrarily slowly. Then we have*

$$|E_\delta(X,h)| \ll_\delta X^{1/2}\log^2 X \; g(X) \qquad \textit{for every } \delta > 0.$$

Proof.

Let $g(x) \to \infty$ arbitrarily slowly, we subdivide $[X, 2X]$ into $\ll g(X)$ intervals of the type $I_j = [X_j, X_j + Y]$ with

$$X \ll X_j \ll X \qquad \text{and} \qquad Y \ll \frac{X}{g(X)}. \tag{14}$$

Then we have

$$|E_\delta(X,h)|X \ll \int\limits_{E_\delta(X,h)} |h(x)|^2 dx$$

$$\ll_\delta \int\limits_{E_\delta(X,h)} |\Delta(x,h)|^2 dx \ll \sum_j \int\limits_{E_{\delta,j}(X,h)} |\Delta(x,h)|^2 dx,$$

where,

$$E_{\delta,j}(X,h) = \{X_j \le x \le X_j + Y : |\Delta(x,h)| \ge \delta h(x)\}.$$

Writing

$$\xi_j = 2X_j^{-1/2} \tag{15}$$

we have

$$\Delta(x,h(x)) - (\psi(x+\xi_j x) - \psi(x) - \xi_j x) \ll \frac{Y}{\sqrt{X}} = o(X^{1/2})$$

uniformly in j and $x \in I_j$, which implies

$$\sum_j \int\limits_{E_{\delta,j}(X,h)} |\Delta(x,h)|^2 dx = \sum_j \int\limits_{E_{\delta,j}(X,h)} |\Delta(x,\xi_j x)|^2 dx + o(X\;|E_\delta(X,h)|).$$

Then we have

$$|E_\delta(X,h)| \ll_\delta \frac{\sum_j \int_X^{2X} |\Delta(x,\xi_j x)|^2 dx}{X}.$$

Recalling that Saffari and Vaughan [9] proved, under RH, that

$$\int_X^{2X} |\psi(x+\theta x) - \psi(x) - \theta x|^2 dx \ll \theta X^2 (\log\frac{2}{\theta})^2, \tag{16}$$

uniformly in $0 < \theta \le 1$, we get

$$|E_\delta(X,h)| \ll_\delta \frac{\sum_j X^2 \xi_j \log^2 X}{X},$$

and then Lemma 2 follows from (14) and (15).

3. - PROOF OF THEOREMS 1 AND 2.

We define the set

$$A_\delta(N) = \{\sqrt{N} \le n \le \sqrt{2N} : |\psi((n+1)^2) - \psi(n^2) - (2n+1)| \ge \delta(2n+1)\},$$

that contains the exceptions, if any, to the expected asymptotic formula for the number of primes in intervals of type $[n^2,(n+1)^2]$ in $[N,2N]$. Moreover define

$$\eta_\delta = \inf\{\xi \ge 0 : |A_\delta(N)| \ll_\delta N^\xi\}$$

and

$$\eta = \sup_{\delta>0} \eta_\delta.$$

From these definitions we have that the asymptotic formula (1) holds with $O(N^{\eta+\varepsilon})$ exceptions, for every $\varepsilon > 0$.

The main step of the proofs is to prove that for every $\varepsilon > 0$ we obtain

$$\eta \le \mu - \frac{1}{2} + \varepsilon. \tag{17}$$

In order to prove (17) we let $n \in A_\delta(N)$, $x = n^2 \in [N,2N]$ and $h(x) = 2\sqrt{x}+1$. From the definition of the set $A_\delta(N)$ we get

$$|\psi((n+1)^2) - \psi(n^2) - (2n+1)| \ge \delta(2n+1),$$

and then

$$|\psi(x+h(x)) - \psi(x) - h(x)| \ge \delta h(x),$$

which implies $x \in E_\delta(N,h)$. Using $i)$ of Theorem 1 of [2], with $\delta' = \delta/2$, we obtain that exists an effective constant c such that

$$[x, x+ch(x)] \cap [N,2N] \subset E_{\delta/2}(N,h).$$

Let $m \in A_\delta(N)$, $m > n$. As before we can define $y = m^2 \in [N,2N]$ such that

$$[y, y+ch(y)] \cap [N,2N] \subset E_{\delta/2}(N,h).$$

Choosing $c < 1$ we get

$$y - x = m^2 - n^2 \ge (n+1)^2 - n^2 = 2n+1 = 2\sqrt{x}+1 > ch(x).$$

Thus we have

$$[x, x+ch(x)] \cap [y, y+ch(y)] = \emptyset,$$

and then

$$|A_\delta(N)| \ll_\delta \frac{|E_{\delta/2}(N,h)|}{\sqrt{N}} \ll \frac{N^{\mu+\varepsilon}}{\sqrt{N}} \ll N^{\mu-1/2+\varepsilon}, \tag{18}$$

for every $\delta > 0$ and $\varepsilon > 0$.

Hence we have

$$\eta_\delta \leq \mu - \frac{1}{2} + \varepsilon \quad \text{for all} \quad \delta > 0,$$

and then (17) follows.

Using (17) and Lemma 1 we can prove Theorem 1.

Using (18) and Lemma 2 we get

$$|A_\delta(N)| \ll_\delta \frac{|E_{\delta/2}(N,h)|}{\sqrt{N}} \ll_\delta g(N)\log^2 N,$$

for all $\delta > 0$ and $g(x) \to \infty$ arbitrarily slowly and then Theorem 2 follows.

3. - PROOF OF THEOREM 3.

In order to prove Theorem 3 we assume that (1) does not hold. Then there exists $\delta_0 > 0$ and a sequence $x_j \to \infty$ with $|\Delta(x_j,h)| \geq \delta_0 h(x_j)$ and

$$h(x) = 2\sqrt{x} + 1. \tag{19}$$

For x_j sufficiently large, choose $\delta' = \delta_0/2$ in $i)$ of Theorem 1 of [2]. Hence

$$|\Delta(x,h)| \geq \frac{\delta_0}{2} h(x) \geq \frac{\delta_0}{2}\sqrt{x_j} \quad \text{for} \quad x_j \leq x \leq x_j + \frac{\delta_0}{20}\sqrt{x_j}.$$

From our assumption then we get

$$x_j^{3/2} \ll \int_{x_j}^{x_j+Y} |\Delta(x,h)|^2 dx, \tag{20}$$

where

$$Y = \frac{\delta_0}{20}\sqrt{x_j}. \tag{21}$$

From (19) we see that

$$h(x) = h(x_j) + O(1) \quad \text{uniformly for} \quad x_j \leq x \leq x_j + Y$$

and hence

$$\int_{x_j}^{x_j+Y} |\Delta(x,h)|^2 dx = \int_{x_j}^{x_j+Y} |\psi(x+h(x_j)) - \psi(x) - h(x_j)|^2 dx + O(x_j^{1/2}). \tag{22}$$

From the definition of the functions $\psi(x)$ and $\theta(x)$ is not difficult to prove

$$\int_{x_j}^{x_j+Y} |\psi(x+h(x_j)) - \psi(x) - h(x_j)|^2 dx =$$

$$\int_{x_j}^{x_j+Y} |\theta(x+h(x_j)) - \theta(x) - h(x_j)|^2 dx + O(x_j^{1/2}\log^2 x_j). \tag{23}$$

From (20), (22) and (23) we get

$$x_j^{3/2} \ll \int\limits_{x_j}^{x_j+Y} |\theta(x+h(x_j)) - \theta(x) - h(x_j)|^2 dx = J(x_j+Y, h(x_j)) - J(x_j, h(x_j)).$$

Assuming Conjecture 3 we obtain that for j sufficiently large we get

$$x_j^{3/2} \ll J(x_j+Y, h(x_j)) - J(x_j, h(x_j)) = o(x_j^{3/2})$$

a contradiction, and then Theorem 3 follows.

References

[1] D. Bazzanella. *Primes in almost all short intervals II. to appear*, 1998.

[2] D. Bazzanella and A. Perelli. *The exceptional set for the number of primes in short intervals. to appear*, 1998.

[3] H. Davenport. *Multiplicative Number Theory*, volume GTM 74. Springer - Verlag, 1980. second edition.

[4] D.A. Goldston. *Linnik's theorem on Goldbach numbers in short intervals. Glasgow Math. J.*, 32 (1990), 285–297.

[5] D.A. Goldston and M.L. Montgomery. *Pair correlation of zero and primes in short intervals. Analytic Number Theory and Dioph. Prob., ed. by A.C. Adolphson et al.*, pages 183–203, Birkhäuser Verlag 1987.

[6] D.R. Heath-Brown. *The difference between consecutive primes II. J. London Math. Soc. (2)*, 19 (1979), 207–220.

[7] D.R. Heath-Brown. *Zero density estimates for the Riemann zeta-function and Dirichlet L-function. J. London Math. Soc. (2)*, 19 (1979), 221–232.

[8] M. Jutila. *Zero-Density estimates for L-functions. Acta Arith.*, 32 (1977), 52–62.

[9] B. Saffari and R. C. Vaughan. *On the fractional parts of x/n and related sequences II. Ann. Inst. Fourier, Grenoble*, 27 (1977), 1-30.

[10] A. Selberg. *On the normal density of primes in small intervals, and the difference between consecutive primes. Arch. Math. Naturvid.*, 47 (1943), 87–105.

Danilo Bazzanella
Dipartimento di Matematica
Politecnico di Torino
Corso Duca degli Abruzzi, 24
10129 Torino, Italy
e-mail : bazzanella@polito.it